DOCUMENT RESUME

ED 047 659 24 HE 002 044

AUTHOR	Yinger, J. Milton; And Others
TITLE	Middle Start; Supportive Interventions for Higher Education among Students of Disadvantaged Backgrounds. Final Report.
INSTITUTION	Oberlin Coll., Ohio.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-5-0703
PUB DATE	24 Nov 70
GRANT	OEG-6-10-255
NOTE	186p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	*Academic Performance, *Compensatory Education Programs, Culturally Disadvantaged, *Disadvantaged Youth, *High School Students, *Special Programs
IDENTIFIERS	*Oberlin College

ABSTRACT

Since 1964, Oberlin College has conducted a "special opportunity program" for post-seventh graders from deprived backgrounds. Each year approximately 65 students have been brought to the campus for an intensive period of educational, artistic, and recreational experience. This report concerns the post-program educational attainments of the first three summer groups, a total of 195 students. Schools in St. Louis, Mo., and Cleveland, Lorain, Elyria, and Oberlin, Ohio were asked to nominate twice as many individuals from culturally deprived or poverty backgrounds who had some likelihood of success as the program was able to accept. From each school's list pairs were matched as closely as possible, and one from each pair was randomly selected to participate in the program. This study tried to determine whether the summer experience, plus the follow-up contacts with the Special Opportunity staff, significantly improved the likelihood of a person staying in school, attaining good grades, and for the 1964 group, of entering college. Significance was measured against the performance of the randomly selected controls. This report also describes the method that was used for taking account of the quality of the match and examines the importance of networks of "significant others." (AF)

BR5-0703
PA 24

HE

ED047659

FINAL REPORT
Project No. 5-0703
Grant No. OE 6-10-255

MIDDLE START

Supportive Interventions for Higher Education among Students of Disadvantaged Backgrounds

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

J. Milton Yinger
Kiyoshi Ikeda
Frank Laycock
of
Oberlin College
Oberlin, Ohio

November 24, 1970

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

HE002 044

FINAL REPORT
Project No. 5-0703
Grant No. OE 6-10-255

MIDDLE START
Supportive Interventions for Higher Education among Students of Disadvantaged Backgrounds

J. Milton Yinger
Kiyoshi Ikeda
Frank Laycock
of
Oberlin College
Oberlin, Ohio

The research reported herein was performed pursuant to a grant from the Office of Education, U.S. Department of Health, Education, and Welfare. Grantees undertaking such projects under Government sponsorship are encouraged to express freely their professional judgments in conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. Department of
Health, Education, and Welfare

Office of Education
Bureau of Research

TABLE OF CONTENTS

Page
Summary iv
Chapter One. Theoretical Perspective 1
The Field Theoretical Approach 8
Recent Literature on Educational Intervention 14

Chapter Two. Methodological Perspective 21
The Origins of the Oberlin College Special Opportunities Program 21
Post-Summer Follow-Up 24
Selection Procedures 28
The Research Design 29
Operational Measures of Initial and Long-Term Program Impact 33
Summary of Methodological Procedures 38

Chapter Three. Background of the Middle Starters 39
Size and General Composition of the Experimental Group of Pupils 40
Bases for Choosing Experimental Pupils 43
Home Background 46

Chapter Four. Testing an Hypothesis by Use of Matched Pairs . 52
The Role of Matching in Sociological Experiments 56
"Ideal" versus "Directional" Matching 67
Closeness of Match between Experimental-Control Pairs 74
Postscript 84

Chapter Five. Significant Others and Educational Attainment . 86
The Participation of Administrative-Legal Functionaries as Significant Others 93
Findings in Connection with Significant Others 98
Access to Significant Others 101

Chapter Six. Conclusion .106

Appendix
A I. A Note on Hunt and Hardt's Findings on the Effects of Upward Bound Programs on the Academic Achievements of Negro Students A-1
A II. Crucial Decisions in Preparing Telented Urban Youths for College A-13
A III.Reforms as Experiments and Experiments as Reforms . . . A-32
B I. Suggestions for Selection Procedures for Summer Special Educational Opportunities Program B-1
B II. Other Instruments and Documents B-7

Tables

Page

Table I-1 Departure from Expected Percentage of Males Graduating from College among a Sample of Wisconsin Youths 12
Table II-1 Types of Matching 33
Table II-2 Archival Information 34
Table II-3 Interviews 36
Table II-4 Questionnaire and Counselor Returns 37
Table III-1 Distribution of Pupils by Year of Participation, Sex, and Race 40
Table III-2 Cities from which Pupils Came 41
Table III-3 Distribution of Ages at time of Entering Program 42
Table III-4 Tested Intelligence Levels at Time of Entry 44
Table III-5 Mean Grade-Level Scores on Achievement Tests at Time of Entry 45
Table III-6 Mean Class Marks at Time of Entry 45
Table III-7 Siblings in the Home 46
Table III-8 Father's Status in the Home 47
Table III-9 Mother's Status in the Home 47
Table III-10 Education of Father and Mother 48
Table III-11 Father's Occupational Skill Level 49
Table III-12 Mother's Occupational Skill Level 50
Table III-13 Mother's Occupational Status 50
Table III-14 Parents' Religious Preference 51
Table IV-1 Distribution of Matching Scores: Ideal Index 76
Table IV-2 Distribution of Matching Scores: Directional Index 77
Table IV-3 Comparison of Post-Program Achievement Levels of Matched Pairs 79
Table IV-4 Shift of Advantage Score as a Result of the Program 80
Table IV-5 Pre-Program to Post-Program Gain or Loss in Academic Performance between Matched Pairs 81
Table V-1 Relative Standing on the Index of Educational Achievement by Initial Level of Educational Advantage and Program Participation 100
Table V-2 Mean Differences and Pair-by-Pair Differences in Access to Significant Others by Program Participation and by Resource/Achievement Level 102
Table V-3 Higher Access to Significant Others among Experimentals by Resource/Achievement Levels 104

Charts

Page

Chart II-1 A Patchy Institutional Cycle Design 31

Figures

Figure IV-1 Illustrative Correlations between "Matching Index" Scores and "Differential Achievement" Scores among Matched Pairs 66
Figure IV-2 Possible Correlations between Differences in Predisposition and Differences in Academic Success between Matched Pairs 71
Figure IV-3 Possible Curvilinear Correlations between Predispositions to Academic Success and Actual Success among Matched Pairs 71
Figure IV-4 Possible Types of Advantage, Using Ideal and Directional Matching Processes 72

SUMMARY OF THE MIDDLE START PROGRAM

In the summer of 1964, Oberlin College began a "special opportunity program" for post-seventh graders from deprived backgrounds. Each summer since then, approximately 65 students have been brought to the campus for an intensive period of educational, artistic, and recreational experience. This report concerns the post-program educational attainments of the participants in the first three summer groups-- a total of 195 students.

Children were selected from schools in St. Louis, Missouri, and from Cleveland, Lorain, Elyria, and Oberlin, Ohio. Two-thirds of the participants were black and two-thirds were boys. Each school was asked to nominate to us twice as many pupils as we were able to take.

We asked them to nominate individuals who had some likelihood of success despite their "culturally deprived" or poverty backgrounds. If school grades or test scores were not reasonably good, we wanted at least to know that some teachers recognized a spark of initiative. Thus this research deals with students from deprived backgrounds who were "visible" in their schools; but there is a wide range in their pre-program performance. From each school list, pairs were matched as closely as possible, and one from each pair was randomly selected to participate in the program. The study seeks to determine whether the summer experience, plus the follow-up contacts with the Special Opportunity staff, significantly improved the likelihood of a person staying in school, attaining good grades, being selected for a special academic school, and--for the 1964 group (which has already finished high school) entering college. Significance is measured against the performances of the randomly selected controls. A new method for taking account of the quality of the match is described and employed in assessing the affects of the program. The importance of networks of "significant others" is also examined.

In a series of measures, it is shown that the experimental group makes a significantly higher rating than the control group. This advantage is not eliminated when the quality of the match is taken into account. That is, the higher academic attainments of the experimental students were not preordained because of pre-program advantages. Not all the students in the program gained, to be sure, or even held their own in relationship with their matched partner in the control group. With pre-program advantages taken into account, 43 per cent of the experimental children gained, 36 per cent held their own, and 21 per cent lost, relative to their partners.

We were slow in getting our follow-up program under way, after the initial summer. There were also periods during which it lagged, due to staff changes and modest financing. It did not include, as we would have liked to have done, a second summer on the campus. In our judgment, a program that included two intensive summer periods and continuous contact with the staff during the school year--through reunions, Saturday "seminars," newsletters, and the like--could significantly lift the educational performance of perhaps two-thirds of educationally deprived "middle starters." The cost would be about $3,000 per child--an investment with an enormous potential rate of social, and even financial, return.

MIDDLE START

by

J. Milton Yinger

Kiyoshi Ikeda

Frank Laycock

Chapter One: THEORETICAL PERSPECTIVE

Explanations of human behavior are subject to waves of style. Today we see some resurgence of the belief that man's capacity for aggression is best understood by studying his biological inheritance and his place in an evolutionary process.[1] There are some who emphasize the biological element in intelligence, with particular attention to individual and racial differences.[2] Others would recognize the impact of experience, but from a classic or modified Freudian perspective see only the earliest years as crucial.[3] Those are the years in which a rich and flexible or a constricting lingual screen is formed; a sense of self and a motivational structure that support intelligence and confidence, or inhibit them, are developed; nutritional and health experience promote development or create a permanent deficit.

---

[1] See, for example, Robert Ardrey, TERRITORIAL IMPERATIVE, Atheneum, 1966; Konrad Lorenz, ON AGGRESSION, Harcourt Brace and World, 1960.

[2] Arthur R. Jensen, "How Much Can We Boost IQ and Scholastic Achievement?" HARVARD EDUCATIONAL REVIEW, Vol. 39, 1969, pp. 1-123; Audrey M. Shuey, THE TESTING OF NEGRO INTELLIGENCE, 2nd ed., Social Science Press, 1966.

[3] Esther P. Edwards, "Kindergarten is Too Late," SATURDAY REVIEW, June 15, 1968, pp. 68 ff. For more analytic and less pessimistic examinations of this question, see the chapters by J. M. Hunt, Carl Breiter, and Celia B. Stendler-Lavatelli in SOCIAL CLASS, RACE, AND PSYCHOLOGICAL DEVELOPMENT, Martin Deutsch, Irwin Katz, and Arthur R. Jensen, eds., Holt, Rinehart, and Winston, 1968, chaps. 8, 9, and 10.

Many beliefs, theories, and policies, on the other hand, emphasize the continuing importance of experience and opportunity. Ceilings on intelligence are set much more fully by available cultural resources than by genes.[4] The federal government supports Head Start programs for three- to five-year-olds and Upward Bound programs for high school students in the expectation that unused capacities can be identified and activated. Colleges and universities throughout the country enroll students who do not meet what they had come to believe were minimum standards, hopefully, if not confidently expecting that educational deficits of the first eighteen years of life can somehow be significantly reduced at the college level.

Those who emphasize the environmental influence on behavior do not, in most instances, deny the importance of constitutional factors and the impact of the first few years. Some psychological behaviorists and sociological structuralists, to be sure, give major attention to external stimuli in their research; but there are few today who overlook, at least in the theoretical statements of their problems, the range of influences on human behavior. Among those who emphasize the external forces, there is, however, an important disagreement:

Some believe that shared normative systems and values--cultures and subcultures--are critical factors. If there are differences in educational aspirations and performances between two groups, for example, these differences are seen as indicators of cultural contrast. This point of view is developed most fully in a way relevant to our research by those who speak of a "culture of poverty."[5]

---

[4] Robert Faris, "Reflections on the Ability Dimension in Human Society," AMERICAN SOCIOLOGICAL REVIEW, Vol. 26, 1961, pp. 835-843; J. M. Hunt, in Deutsch, Katz and Jensen, op. cit., ch. 8.

Other researchers--we shall call them "structuralists," in contrast with the "culturalists"--believe that differences in aspiration and educational performance, along with economic, familial, and other differences, are primarily a result of the structure of opportunities. On the one hand, they emphasize the extent to which values and aspirations are shared throughout the society, across status levels; on the other hand, they stress the wide variation in resources and opportunities. In the area of our interest, if educational plans and performance vary widely among status groups, this is primarily, the structuralists argue, a product of differential opportunities. It is not, to a significant degree, the result of differential talent nor of widely varying cultural supports.[6] If members of disadvantaged groups are given high school counseling sensitive to their needs, "GI Bills" or other financial supports, community colleges within commuting distance, and the like, the gap between their educational aspirations and their expectations is substantially closed.[7]

---

[5] On the "culture of poverty" see, for example, Oscar Lewis, FIVE FAMILIES: MEXICAN CASE STUDIES IN THE CULTURE OF POVERTY, Basic Books, 1959; Michael Harrington, THE OTHER AMERICA: POVERTY IN THE UNITED STATES, Macmillan, 1962; Herbert Gans, THE URBAN VILLAGERS, Macmillan, 1962; Walter B. Miller, "Lower Class Culture as a Generating Milieu of Gang Delinquency," JOURNAL OF SOCIAL ISSUES, Vol. 14, 1958, pp. 5-19. For critiques, see Charles Valentine, CULTURE AND POVERTY, Univ. of Chicago Press, 1968; Jack Roach and Orville Gursslin, "An Evaluation of the Concept 'Culture of Poverty,'" SOCIAL FORCES, March, 1967, pp. 383-392.

[6] On the effects of school networks, see, for example, James S. Coleman, EQUALITY OF EDUCATIONAL OPPORTUNITY, U. S. Government Printing Office, 1966; David Hargreaves, SOCIAL RELATIONS IN A SECONDARY SCHOOL, Humanities Press, 1967; Robert Rosenthal and Lenore Jacobson, PYGMALION IN THE CLASSROOM, Holt, Rinehart, and Winston, 1968. Many of these issues had been pointed to in early studies. See, for example, W. Lloyd Warner, et al., WHO SHALL BE EDUCATED, 1944.

In the face of these diverse beliefs and theories, we call for a more pragmatic and open-ended approach. Let us find by careful research those elements in the cycle of causation that are most subject to control. It is much more important, in our judgment, to develop the enormous store of unused capacity in all human beings than to debate the range of variation in ability set by existing circumstances.

That does not mean we started our research without judgments on questions related to the sources of talent. In some measure we share each of the perspectives referred to above. Each must be qualified in important ways, however, and then all must be brought into one theoretical system that takes full account of their interdependence. We shall comment briefly on each of these points.

It appears to us that there is some range in human capacities influenced by inheritance. Two qualifications are needed, however, to put this observation into perspective:

1) There is no evidence that the range of inherited capacity varies significantly among races. Variation in group levels of measured

---

[7] On the gap between aspirations and expectations, see, for example, Wan Sang Han, "Two Conflicting Themes: Common Values Versus Class Differential Values," AMERICAN SOCIOLOGICAL REVIEW, October, 1969, pp. 679-690; R. N. Stephenson, "Mobility Orientation and Stratification of 1,000 Ninth Graders," AMERICAN SOCIOLOGICAL REVIEW, April, 1957, pp. 204-212; Hyman Rodman, "The Lower Class Value Stretch," SOCIAL FORCES, December, 1963, pp. 205-215.

intelligence reflect the nature of the skills tested, the measuring instruments, and patterns of experience and opportunity.[8]

2) The ceiling on intelligence is set much more by socially shared knowledge and methods of training than by inheritance. The average man today can understand aspects of the natural world that baffled the geniuses of early centuries, because today, as Newton said, we can stand on the shoulders of giants. It is therefore far more important to understand and improve methods of training that lift the base of socially shared knowledge than to emphasize or base policy on the range of inherited differences.

We also recognize the importance of early years of life and the great losses attributable to lack of stimulation, ego-strength, and nutrition as a result of early deprivation. This point must also be qualified, however, in two ways:

1) The importance of early years can be exaggerated, because those who are deprived in infancy are usually those who are also deprived later. We cannot attribute lesser motivation or talent to childhood experiences until we control for the reenforcement effects of later experiences. (It is our hope that the research reported here will make some contribution to this problem.)

---

8. As the geneticist, Hirsch, has said in reaction to Jensen's recent conclusion that inherited factors account for the bulk of intelligence differences, a heritability estimate is a piece of knowledge that is "both deceptive and trivial." ". . . High or low heritability tells us absolutely nothing about how a given individual might have developed under conditions different from those in which he actually did develop." (London TIMES EDUCATIONAL SUPPLEMENT, July 24, 1970, p. 9.)

2) The importance of the first few years is not intrinsic to them, it is a function of our present level of knowledge. We may learn to overcome nutritional deficits, as well as to prevent them, and to reverse or redirect motivational systems.

The more recent emphasis on subcultural and cultural influences has, in our judgment, added a needed dimension to the study of factors involved in educational processes. Easy assumptions about the uniformly supportive aspects of American values with regard to education clearly need to be set aside in favor of careful study of the range of values. Here again, however, the contribution of a particular approach--in this instance, an emphasis on cultural variation--is greatest if its own limitations are fully recognized. In connection with the study of cultural factors in education, we call attention in particular to two qualifications:

1) Culture is not best understood simply as an independent variable, which is operative in a given environment as a causal factor, separately or in conjunction with other factors. Culture is a process; it develops and changes. From one perspective it can be read as an adjustment to the circumstances of a group. Seen in this light, the cultural process is understood only by studying it in the lives of particular people in particular circumstances.[9]

2) Related to this, we would emphasize the range of cultural themes in complex societies. There is, in the United States, a wide variety of cultural influences that bear on education, only some of

---

[9] See Richard Ball, "A Poverty Case; The Analgesic Subculture of the Southern Appalachians," AMERICAN SOCIOLOGICAL REVIEW, December, 1968, pp. 885-895.

which are given emphasis in particular contexts. Certain cultural elements take on salience under the life conditions faced by given individuals and groups.[10]

The networks of interaction into which an individual is bound are of great importance in his educational life, as they are in his economic, political, and religious life. Segments of the social structure, as we are using the term, range from small-scale, personal encounters, to reference groups and significant others in a wider circle, to the impersonal resources of local and national institutions. Together they make up the opportunity system available to an individual. Regardless of his individual capacities and in spite of a strong cultural value placed on education, the level of educational attainment will be low in the absence of structural supports. This proposition must also be qualified:

1) There are individual differences in response to the same structure of opportunities. This means that changes in opportunity alone do not necessarily produce the behavioral changes which were the goal. Interaction effects, as we shall continually emphasize, are crucial.

2) Cultures also vary in the degree to which those who share them are prepared to take advantage of increased opportunity. We think it unwise to explain behavioral differences between classes by reference to culture alone, as we have noted. It is equally unwise, however, to overlook the fact that societies, and groups within societies, evaluate

---

[10] An analogous point is made, with reference to delinquency, by David Matza and Gresham Sykes, "Juvenile Delinquency and Subterranean Values," AMERICAN SOCIOLOGICAL REVIEW, October, 1961, pp. 712-719.

education differently. Persons of equal talent, facing equal opportunities, will respond differently if the cultural supports for a given activity vary.

THE FIELD THEORETICAL APPROACH

A theory adequate to our task must combine attention to the biological, psychological, cultural, and structural influences on education. We shall not attempt to state such a theory here,[11] but will simply illustrate how it might redefine questions related to education and educational interventions. One of the postulates of field theory is that all four influences on behavior must be taken into account: behavior is a product of their interaction. If any one of the four influences is lacking, with respect to a given activity, therefore, that activity cannot occur.[12] Whether by a multiplicative model, an additive formulation, or emphasis on a combination of "interactive" and additive factors, this theoretical perspective suggests that there are systematic, contingent effects on educational conduct stemming from each source or factor. Biological, cultural, structural, and personal influence combine to produce a given outcome. That outcome would not be expected to occur without the joint presence of all the factors, although the weights and combination of weights to be assigned each factor under various conditions remain open to empirical verification.[13]

---

[11] For an attempt to formulate such a theory, see J. Milton Yinger, TOWARD A FIELD THEORY OF BEHAVIOR, McGraw-Hill, 1965. On the concept of "interaction," see John Atkinson, "Motivational Determinants of Risk-Taking Behavior," PSYCHOLOGICAL REVIEW, Nov., 1957, pp. 359-372; Hubert Blalock, Jr., "Theory Building and the Statistical Concept of Interaction," AMERICAN SOCIOLOGICAL REVIEW, June, 1965, pp. 374-380; and Bruce Eckland, "Genetics and Sociology: A Reconsideration," AMERICAN SOCIOLOGICAL REVIEW, April, 1967, pp. 173-194. For empirical studies that emphasize interaction and multivariate analysis of educational behavior, see Ralph H. Turner, THE SOCIAL CONTEXT OF AMBITION, Chandler Publishing Co., 1964; William Sewell, Archibald Haller and Alejandro Portes, "The Educational and Early Occupational Attainment Process, AMERICAN SOCIOLOGICAL REVIEW, February, 1969, pp. 82-92.

[12] A corollary to this is that efforts to improve educational performance are most likely to be successful when several factors are influenced.

The argument can be put in simple mathematical terms, using purely fictional numbers. Assume that a given level of educational performance is a function of inherited capacity, of learned skills, motives, and other tendencies, of cultural definitions of good or appropriate behavior, and of the level of structural opportunities. Assume further that each of these factors can be given a support "score" ranging from 0 to 10. Compare two individuals of similar capacity (5), but with differences in the other three factors. The educational outcomes, the product of all four factors, are sharply different. In a strictly multiplicative model, they would be as follows:

	Inherited capacity	Learned motives and tendencies	Cultural support	Structural support	Product score
Individual A	5	5	5	5	625
Individual B	5	2	2	2	40

---

[13] In systematic, empirical studies, actual effects may reveal both main effects and joint, interactive effects in diverse ways, as shown by the use of a wide variety of statistical tools. See the systematic efforts both to measure and to evaluate the effects of different variables upon school-test performance (attitudinal characteristics of student, characteristics of the student's peer group, environmental characteristics, teacher quality, and school resources) in the EQUALITY OF OPPORTUNITY report. Coleman's discussion of efforts to partition the weights from selected variables and to determine main effects and joint effects (in a non-regression analysis sense) indicates the search for order in these sorts of data. James S. Coleman, "Reply to Cain and Watts," AMERICAN SOCIOLOGICAL REVIEW, April, 1970, pp. 242-252. Also Glen G. Cain and Harold W. Watts, "Problems in Making Policy Inferences from the Coleman Report," AMERICAN SOCIOLOGICAL REVIEW, April, 1970, pp. 228-241; Dennis J. Aigner, "A Comment on Problems in Making Inferences from the Coleman Report," AMERICAN SOCIOLOGICAL REVIEW, April, 1970, pp. 249-252.

If this statement of the problem is correct, it is clear that Individual B, who is educationally deprived, cannot be significantly helped by modifying one factor only. Suppose that he has an experience which strongly increases the motives, skills, and tendencies that he brings to an educational situation (imagine a "score" of 8), but that cultural and structural supports remain low. In a strictly multiplicative model, there is some improvement (5 x 8 x 2 x 2 = 160); but he still falls seriously behind Individual A, who has equal capacity and poorer motivation, but substantially stronger cultural and structural supports.

In another sense, this simple mathematical way of stating the situation indicates that "no chain is stronger than its weakest link." The largest product score of four factors which total 20 is attained with a 5x5x5x5 combination, the lowest (if all zeros are excluded), by any combination of two nines and two ones. If our interpretation is correct, wise policy deals with all possible factors in the educational complex. Theories that emphasize one factor are not only less powerful as analytic tools, they are less useful guides to community action than multi-factor theories. An extraordinarily low score in one factor is likely to mean that "compensatory efforts" in related factors can have only moderate effect.

One way of indexing the relative influence of "opportunity" effects and of "capacity" effects on educational outcomes is to re-examine Sewell's data on a sample of Wisconsin males.[14] In a series of papers, he and his colleagues have sought to identify the factors which determine who is likely to graduate from college. "Opportunity" effects stem from advantaged or less advantaged location on the socio-economic status levels within the society at large. Coming from a higher socio-economic status (SES) as indexed by Sewell, the upper middle class child will have grown up in a more "advantaged" location in the social structure and be exposed to cultural influences which support certain motives and skills necessary for higher educational attainment.[15] An individual of similar "capacity" (as indexed by an IQ measure, crude as it may be for equating capacities or potential), but with parents from lower SES settings, is more likely to be isolated both from opportunities to participate in valued activities and to acquire the kind of culturally shared knowledge found in upper SES settings. Such limitations can be expected to restrict chances for graduation from college, regardless of equality of "capacity."

---

[14] See William H. Sewell and Vimal P. Shah, "Socioeconomic Status Intelligence, and the Attainment of Higher Education," SOCIOLOGY OF EDUCATION, Winter, 1967, pp. 1-23.

If we take the percentage of all males graduating from college as the most likely outcome for the sample of Wisconsin males (21.8%), and compare those of equal capacity against this standard, we would expect the individual from a higher SES setting to go beyond the norm, given his higher chances for learned skills, appropriate motives, and related opportunities and cultural supports. The converse should be true for a male of equal capacity but of lower SES background. As reported in Table I, the relative advantage of a higher SES male completing college work is systematically higher at every level of capacity or IQ. At the

---

15. This statement of the problem collapses the four-variable model into a two-variable model. It is extremely difficult to separate cultural from structural influences, and inherited from learned tendencies, in empirical studies. We will generally use a two-variable model.

Table I

Departure from Expected Percentage* of Males Graduating From College Among a Sample of Wisconsin Youths[16]

Socioeconomic Status Levels	Intelligence Levels				
	Low	Lower Middle	Upper Middle	High	Total
Low	001	036	050	092	034
Lower Middle	010	034	077	158	065
Upper Middle	020	045	112	214	100
High	048	107	177	293	193
Total	014	053	110	217	100 (21.8%)
	(1,070)	(1,100)	(1,084)	(1,133)	(4,386)

*Expected percentage represents the percentage of all males in the total sample who graduated from college. The percentage is 21.8% of all males. This expected value is divided into the obtained percentages for each condition or cell of SES x Intelligence Level to index the departure from the expected value. This value is further multiplied by 100 to round out the values into whole numbers.

highest level of IQ, the child from the lowest SES levels falls just below parity (indexed by a score of 100), with a score of 092. The child at the highest level of both SES and of capacity exceeds parity by an index score of 293, 2.93 times that expected for the population as a whole. In general, the importance of capacity or potential for college performance is recognized by educators. This is revealed in Sewell's analysis and will be shown by the data we present. On the other hand, the loss in educated talent due to lack of opportunities and cultural exposure to necessary skills and motives is also demonstrated very clearly. The ability of a higher SES parent and related adults to support college plans of their children to completion is clearly shown. The chances for completion systematically favor the higher SES child of equal, measured talent over the lower SES child.[17]

---

[16] Adapted from Sewell and Shah, op. cit., p. 15.

Against such a theoretical background, it is clear there may be serious loss of potential among the highly talented youths in lower SES settings. With capacity as a constant, we have sought to determine how motives and skills and socio-cultural opportunities might be increased. We sought out youths with relatively high potential for college among neighborhood and school settings likely to be low in opportunities for college-bound activities. Our experimental summer program and related follow-up activities attempted to increase chances for participation in school programs and also to support educational, cultural, and social activities with these youths and their families. At the same time, given the ways in which social and cultural mobility can create a sense of "psychological marginality," we have sought to develop curricular and pedagogical arrangements which could provide transitional and adaptive responses to possibly divergent socio-cultural demands.[18] We have sought, in short, to determine whether students with some potential for college, exposed to a "Middle Start" program of greater opportunities and college-bound activities along with related efforts to provide transitional experiences for upward mobility, could become college-bound, both in active planning for college and in attending college.

---

[17] Sewell and Shah find that among males, IQ contributes slightly more than SES to graduation from college, op. cit., pp. 17-18. What is important is the joint, contingent presence of both SES and IQ explain the obtained differences.

[18] See Turner, op. cit., pp. 207-210. See also Robert A. Ellis and W. Clayton Lane, "Social Mobility and Social Isolation: A Test of Sorokin's Dissociative Hypothesis," ASR, April, 1967, pp. 237-246, on the effects of recruiting students for college from different social strata and ethnic-racial backgrounds on both the students and the institutions involved.

RECENT LITERATURE ON EDUCATIONAL INTERVENTION

How do recent studies of educational intervention measure up, as guides to theory and action, when judged by a multi-variate field approach? We shall not attempt to review the extensive literature now available; but will refer to some of the recent work to illustrate what we think is the prevailing--and favorable--trend toward more attention to interaction. This trend has not always been supported by clear formulations of a guiding theory; nor has research generally been of an experimental variety that permits firm conclusions. Nevertheless important steps have been taken in the effort to understand the sources of educational performance and to support those children who, in one way or another, are "deprived."

The many efforts today to help minority children have a very short history. The most useful review of those efforts to date is Gordon and Wilkerson's compilation of current school programs for the disadvantaged.[19] It gives brief descriptions for each organized effort in the United States that the authors were able to locate by a massive questionnaire mailing. Nearly all of the programs arose during the 1960's. The accompanying text examines them and their background in a series of pithy chapters.

---

[19] Edmund W. Gordon and Doxey A. Wilderson, COMPENSATORY EDUCATION FOR THE DISADVANTAGED, College Entrance Examination Board, 1966, p. 23 ff.

Gordon and Wilkerson point back to Binet's scientific interest in children whose intelligence he thought could be trained to improve weaknesses. They decry what happened afterward to the test that Binet developed to identify these children. They show how specialists shifted their attention from systematic study of how to train minds, into quantitative methods of classifying intellectual status. This work was often scientific, some of it providing rigorous models for research; but it gave little attention to procedures for improving mental development. Gordon and Wilkerson think that recent efforts to rehabilitate ghetto children's intellectual background would be succeeding better if there had been more work over the past decades on exactly how strengths and weaknesses in the mind arise, and on how to stimulate or correct them.

They cite a second precursor to crash programs for the disadvantaged: the humanitarian-cum-scientific efforts of Maria Montessori with Italian slum children. She was contemporary to Binet, and her work has been curiously resurrected, but in behalf of well-prepared middle-class children. Nonetheless, Montessori developed specific--highly specific--methods of developing mental alertness in the very young. The doctrinaire rigidity of most Montessori disciples ever since the early twentieth century, unfortunately, did not encourage either flexible change or systematic analysis. Thus her techniques have not been as useful to current workers as they might have become.

From our point of view, the most valuable contribution that Gordon and Wilkerson make is their emphasis upon the "interactional" view of how children grow. This view runs sharply counter to two others that command much attention today. According to one, a great many programs in schools assume that a backward or unproductive child suffers a defect.

Improvement means remedying this defect, once it has been carefully located. Any subnormal performance, then, merely needs specific attention to the locus of poor performance. According to the other view, any child who is below par in school is the victim, not of some internal deficit of his own, but of a hostile or bare environment. He must be given experiences to make up for the things his family and neighborhood don't provide, experiences that more favored children absorb all about them as they grow to school age. The first view has been prevalent in the past; the second is fashionable today. It is the thesis of our research, as it is of Gordon and Wilkerson, that neither is sufficient. The organism does not generally come first, with some built-in defect that hampers mental growth (although a very small proportion of children do, of course, have inherited or congenital handicaps). Nor is the environment able to work its effects directly, so that a rich and stimulating home accounts for rapid development, or a spare or frightening one for apathy. Instead, the field view assumes that both the organism and its environment work upon each other dynamically. Hunt's 1961 Intelligence and Experience[20] has been a powerful voice speaking for this idea. Not only does Hunt show that the environment has effect upon an individual depending upon whether prime time has arrived (Havighurst's "teachable moment"). He is most persuasive in showing that individuals, as they respond to their environment, may be sufficiently changed that they in turn are able to act upon and change it. Thus intelligence, far from being fixed (the constant IQ assumption, following Binet), responds to its surroundings by changing its own

---

[20] J. McV. Hunt, INTELLIGENCE AND EXPERIENCE, Ronald, 1961.

"traits." So the straightforward hereditarian or "traitist" and the simple environmentalist are both seeing only part. In particular, they miss the interaction effects.

These theoretical limitations are matched by experimental and action limitations. The major interventions of the past decade have been either for pre-school children or for much older high school dropouts. The present research had very little experience to draw upon for comparison, because the middle school years have attracted virtually no sustained attention. Agreeing with Gordon and Wilkerson that simplistic one-cause-one-effect patterns are too naive, Deutsch[21] points out that both the pre-school and dropout programs rest on unproved assumptions. Head Start (the federal government's premature application of some of Deutsch's own early research into liguistic development) assumes "the earlier the better." Deutsch says that this assumption may well be justifiable, but that hard data are still wanting. Especially is this a vicious assumption if it fosters the rationalization that "after seven or eight years of age all is lost, because formation takes place in infancy." Efforts to salvage dropouts have usually been weak, not only because they have come late, but also because they have been realtively brief and partial. And since they have not been very successful, they, too have bred an unfortunate claim; that "there's nothing to be done, the pattern is already set." Stages of growth not only have their own patterns (each can produce progress, whether in due course or under remediation), they stretch out over connected experiences so that steady progress requires continual attention, continual and adaptable programs. Deutsch is wise when he inveighs constantly against short programs without enough follow-through.

---

[21] Martin Deutsch, "Social Intervention and the Malleability of the Child," Fourth Annual School of Education Lecture, Cornell University, 1965, pp. 6 ff.

The intervention described in our research follows all these lines of advice. It shuns a simplistic search for single causes of failure, preferring the field theoretical view that many activities directed at many facets of development are requisite. It assumes that by junior high age pupils are still malleable. And it refuses to stop after a brief, if very intensive, summer of stimulation, continuing instead to offer support and direction over several critical adolescent years. Unfortunately there was no adequate program available as a model. Upward Bound began a year later. New York's Higher Horizons project offered insufficient evidence. Concurrent Rockefeller-funded interventions aided older pupils, usually late in senior high school. The great gap between pre-school and adolescent programs was indeed evident. With the advantage of hindsight, of course, we can offer many comparisons. Many techniques have been tried by now: special teachers (reading, speech, e.g.), special auxiliary professionals (counselors, physicians), sub-professionals (student teachers and interns, lay specialists). Residential and day-school settings, ingenious teaching methods (teams, homogeneous groups, extended days), community cooperation--these and others have been tried in every part of the country. Now important precursors have been publicized for inspiration and comparison: demonstration guidance projects in Harlem junior high schools, impressive New Mexico efforts in behalf of children from Indian and Mexican backgrounds, NDEA programs for children with latent or unrecognized talent, the Ford Foundation's Great Cities Project. Probably the Rockefeller Panel

Reports, Prospect for America,[22] as much as anything else, pushed interest into intensive effort, by pointing out just how costly it is to waste talent. Its effects, however, have been of direct benefit mainly to high school juniors and seniors. And almost nobody has paid sufficient attention to systematic evaluation, either during or after a program.

Careful evaluation is the crux of this report. Here, too, there was very little to go on, for almost every previous assessment had been *ex post facto*. That is, statistical techniques (often quite sophisticated) examine the characteristics of pupils who were helped, to compare them with carefully chosen peers. The landmark assessment, the so-called Coleman Report of 1966,[23] was not so much a description of intervention as it was an analysis of the national status quo after desegregation. It concluded that the home and neighborhood were powerful forces, so powerful that the school had much less impact. It also pointed to the unequal quality of schools across the country, the systematic differences between white and non-white achievement. But it was a *cross-sectional* study, albeit very well done. *Longitudinal* research would allow the effects of time to show up, and in the same persons year after year. In Wisconsin, Sewell has followed groups of pupils into adulthood since the 1950's with sensitivity and skill. But he is quick to confess that most such work, including his own, has to deal with what happens to take place. "Actually, the effects of schools and of other variables should be determined at least by longitudinal studies and at best by well-designed experiments in which students are assigned to schools at random or, if this is not possible--as it probably is not--there should be prior careful assessments of ability, family background, and other potentially confounding variables so that their effect can be controlled or appraised statistically."[24] There is thus a hierarchy, where cross-sectional studies are least powerful, longitudinal ones a great improvement, but true experiments (with random assignment) best. We know of no intervention into the lives of disadvantaged pupils where candidates for the experimental and control groups have evenhandedly come from the same pool, assigned at random. The comparisons that such an experiment permit are, of course, the ones most suited to generalization beyond the study itself. They are the best base for advice to inquirers who want to know what worked, what didn't, and perhaps even why.

---

[22] Rockefeller Panel Reports, PROSPECT FOR AMERICA, Doubleday, 1961.

[23] James S. Coleman, *et al.*, *op. cit.*

This research, then, is applied to a program of intervention that tried to follow current canons: choose a population that badly needs help, intervene at a still-formative stage, fit the intervention to as many aspects of a pupil's life as possible, continue it over a long enough time to insure dependable effects, and gather all the data possible to permit intelligent shifts during the program and systematic analysis and generalization afterward. The assessment, which went hand in hand, is longitudinal-experimental, applied to experimental and control populations that were paired randomly before the study rather than statistically *ex post facto*.

---

[24] William H. Sewell, Review of Coleman Report, AMERICAN SOCIOLOGICAL REVIEW, June, 1967, pp. 475-479.

Chapter Two: METHODOLOGICAL PERSPECTIVE

The research strategy reported here involves three related efforts to contribute to the systematic evaluation and assessment of educational programs. First, we seek to contribute to knowledge of effects of age by selecting students in the "middle years" of their schooling, without a priori assumptions about rigidity of response to new educational programs. Second, we develop a true experimental field study, involving both random assignment and pre-matching of individuals to increase the interpretability of program effects. Third, we make a major effort to follow our participants through several years (six years for those in the earliest group) in order to get knowledge of the long-run effects of the program.

THE ORIGINS OF THE OBERLIN COLLEGE SPECIAL OPPORTUNITIES PROGRAM

Methodological aspects of the study will be better understood against the background of the origins of the project. In 1963, Oberlin College administrators and faculty members began to discuss the value of institutional efforts to prepare and to recruit talented youths from minority backgrounds to aim for college. The President of Oberlin College, learning of possible Rockefeller Foundation support for such efforts, called together a faculty committee charged with two responsibilities: supervision of a college-level program of financial aid and supplemental educational services, and a pre-college program to prepare secondary school youths for college. To support the pre-college work, the Foundation made a joint award to Oberlin, Princeton, and Dartmouth. Whereas Dartmouth and Princeton--and most other institutions working at pre-college level-- chose to help students in their last year or two of high school, Oberlin decided to begin somewhat earlier. The faculty committee, after full discussion of alternatives, decided to bring a group of junior high pupils (mostly those between seventh and eighth grades) to the campus for an intensive residential summer. The first summer, 1964, 56 pupils came to the campus, from Cleveland, St. Louis, and three communities in the county near Oberlin (Lorain, Elyria, and Oberlin). The children from the St. Louis and Cleveland schools were selected from schools located in neighborhoods which could be defined as "inner city." Two of the junior high schools in Cleveland from which the nominees came were predominantly Negro; one was predominantly Eastern European ethnic. The St. Louis schools--in the Banneker District--were overwhelmingly Negro. Lorain contained a large Spanish-speaking population as well as both Negro and White migrant communities. Oberlin and Elyria were selected because of sizable concentrations of economically and educationally disadvantaged households, both Negro and White. There were 40 students who had completed seventh grade, 16 students who had completed the eighth. Every year since then, only those who have just completed the seventh grade have come.

Each summer program has lasted six weeks. Although costs have compelled us to shift to a day school during the last two summers, the three groups under study were in a residential program. The staff has totaled about 20 during each of the summers, creating a rich staff-pupil ratio to maximize impact. Furthermore, the curriculum has been broad, going well beyond the academic subjects. This breadth was deliberate, because we believed that while a pupil might return home and eventually finish secondary school, his schooling would be seriously weakened. So we worked on other interests and skills during the summer beside those bearing most directly upon academic studies: aesthetics, physical activities, hobbies, personal and social sensitivity. One of the most intensive efforts was during periods of personal counseling, when motivation and opportunity for higher education were explored in detail.

A typical regimen (this one from 1966) was:

7:00: breakfast
8:00-12:00: classes

12:15-1:00: lunch

1:00-2:00: rest and guidance
2:00-3:20: study, guidance, individual help
3:30-5:00: swimming, bowling, modern dance
5:00-6:00: free
6:00-7:00: dinner

7:00-8:30: study
8:30-9:30: individual tutoring, study, recreation
9:30-10:00: snacks

10:00: lights out

On week-ends there was a different schedule, to include field trips (to plays and concerts, museums and factories), small group work, individual tutoring, recreation, church, and free time.

Throughout the six weeks each pupil developed close ties with various other persons: a college student, who worked and lived with a small group of pupils as his special assignment; a counselor, who held several individual and small-group sessions; his teachers, of academic and other subjects; and of course, fellow pupils. Records of various kinds, formal and informal, were kept. The adults met regularly to discuss pupils and the program, and from time to time changes were made to respond to opportunities and needs. It does not require statistical analysis to realize that the summer was intensive and sustained.

POST-SUMMER FOLLOW-UP

Additional effects flow from the follow-up activities developed within the Special Opportunities Program. These activities are designed to enhance and maintain the student's interests and skills in four areas: academic competence, career planning, social and interpersonal competence, and extra-curricular enrichment.

The original program did not provide for sustained follow-up. There was no direct budget allocation, during the first year after the summer program of 1964, for any of the forms of follow-up noted above. In the next year, however, a one-fourth time director was appointed. Part-time secretarial assistance was added to maintain a newsletter, to plan some reunions during the summer, and to provide a book service to enable students to own and develop a personal library. The following year the director was on half-time. Liaison personnel in the schools and among parents were appointed to assist in the development of school and parents' clubs. The following year saw the appointment of a full-time S.O.P. summer director who was also responsible for follow-up programming of the range and quality described below. These details are mentioned because both in fiscal terms and in the range and quality of follow-up services, there are likely to be systematic effects upon the children chosen in the summers of 1964, 1965, and 1966. The 1964 summer participants have received the least follow-up support. Increasing and earlier follow-up efforts apply to the 1965 students and even more to the students who came to the campus in the summer of 1966.

Since the post-summer activities have become a vital part of the program of educational intervention, we shall describe them in some detail. Beginning with procedures and contacts in the summer and going on to ideas and activities developed afterwards, the Oberlin staff coordinated a series of measures to insure continuity. On the summer staff were teachers and counselors from the home schools. In addition to the requirement that they be skillful teachers or counselors, these persons had been hired for the summer because they were on the staff of the pupils' schools and could therefore serve as liaison officers between the summer program and the regular school, between Oberlin and home. Through them, Oberlin kept records, stimulated home contacts and school activities, and watched carefully the academic performance of our participants.

We also made formal contacts with other important officials: administrators and guidance workers in the schools, research and supervisory staff "downtown." These contacts were necessary if continued activity in behalf of our pupils were to be possible and if records were to be forthcoming regularly. In particular, as we reached out into the homes (for visits, interviews, and mutual planning for the pupils' future), we believed that working through local school people already known and trusted would be far better than sending in outsiders from the college alone. Summer staff recommendations were collated and sent to the schools to encourage local school follow-up in course programming and counseling. Support in the schools was devoted mainly to trying to assure that each pupil took the right courses for maximum educational attainment. Typically this meant that he should be assigned to the college-bound track. It meant also that he be helped to succeed in required courses. So counselors were alerted, and tutoring made available. At strategic times, Oberlin representatives would reinforce local people, with reunions, talks, individual sessions. Here the assumption was that most school districts could sharpen their counseling and tracking of the bright but relatively unnoticed children from deprived backgrounds, that many could avail themselves of the expertise of nearby universities. We also supported the formation of parents' clubs, designed to inform participants about school policies and opportunities that could benefit their children, to spread the word about financial aid open to disadvantaged pupils, to keep the goal of further education a live option.

Each group of pupils returned once to Oberlin for a brief summer reunion, where college data (academic and financial in particular) were the special topics. Pupils were encouraged to examine a wide range of colleges and to aim for one that was a good match with their plans and aspirations. More frequently, each pupil also met with the others in his group in his city--to go on trips, to attend concerts or ball games, to see movies or hear talks, or to have a party. Those near Oberlin made their own way to campus occasionally, to visit the college student who had been their summer leader, or to attend a game.

In our follow-up efforts, we have not generally tried to dislodge the child from his original home-school-neighborhood setting. Rather, we have sought to encourage the parents and the S.O.P. liaison staff in the schools to reinforce the child's best efforts to achieve well in school. We have attempted to draw upon and stimulate available local resources and to make them more visible to the children, the parents, and the school staff. Thus, the family remained intact within its setting, and the efforts on behalf of the pupils did not isolate the child or the family from the realities of their past.

For some of our pupils, cumulative experience and policy reviews among follow-up staff encouraged a shift in practice, namely a move to a private, preparatory school. Some of these moves stem directly from program staff efforts; others come from the intense desire of the children and the families, after the summer program, to improve the schooling of the children involved. Neither choice--that of returning the child to settings which might limit his dreams or his efforts or that of placing the child in a new environment that may be strange, novel, and even punishing--is a comfortable one for any of those involved. But further upward mobility, in educational, occupational, and community terms, requires that a child and his family confront variation and strain in social and cultural traditions; no educational program has yet found a way to eliminate the impact of transition and change in social position.[1]

As each pupil came closer to high school graduation (so far, this includes the 1964 and 1965 summer program participants), he was given special attention to make certain that he found out as much as possible and did what must be done to select and apply for college: college characteristics, financial aid, necessary tests, application forms and recommendations were noted. The S.O.P. staff responded to questions about specific problems; they sought scholarships and encouraged pupils to apply for them. At this time the newsletters and visits and reunions that had become a pattern were all used to smooth the high school-post high school transition.

---

[1] See Robert A. Ellis and W. Clayton Lane, "Social Mobility and Social Isolation: A test of Sorokin's Dissociative Hypothesis," AMERICAN SOCIOLOGICAL REVIEW, April, 1967, pp. 237-246. See also Appendix.

Post-Summer Follow-Up and the Study Design. Has our post-design "tinkering" with follow-up plans (efforts which are addressed to fundamental questions about the nature and quality of sustained supports for higher education among "less advantaged" children) made the original design uninterpretable? We believe not. This judgment rests on the fact that much of the summer programming and staffing and the selection procedures remained relatively constant. The major variant in the treatment, as noted above, involved greater effort at sustained follow-up, and such effects are traceable to each cohort of S.O.P. participants in 1964, 1965, and 1966. There is now available a technical treatment on how program changes can be evaluated in experimental design arrangements. Welty's analysis provides the statistical basis for evaluating the differential effects of program changes over time, especially when the decisions are focused in a known direction or set of directions.[2]

SELECTION PROCEDURES

Our over-all pattern has been to invite pupils who had finished the seventh grade, two-thirds of whom were black, and two-thirds boys. We chose a majority of blacks because in the early 1960's, theirs appeared to be the most urgent need. We chose a majority of boys because efforts to help the entire disadvantaged community seemed more strategic when directed at prospective breadwinners. And we dipped down to the seventh grade because we thought such pupils should be more malleable than they would be later on. All of our pupils have come from backgrounds that meet definite specifications of economic hardship, akin to those for Upward Bound programs.[3]

---

[2] Gordon A. Welty, "The Logic of Evaluation," Educational Resources Institute offprint, Washington, D. C., 1969, pp. 7-10, 14-15.

THE RESEARCH DESIGN

Our cooperative research effort involves a study of the impact of the S.O.P. program on the educational careers of the 195 students invited to participate in the first three years of the program. Our final research design is a patchy institutional cycle design (Design #6 or the Fisher After-Only Design) with precision matching before random assignment. The method of precision-matching is described in a later section. As Campbell and Stanley suggest, a true experimental design (A Fisher After-Only Design) is more likely to control invalidity stemming from history, maturation, instrumentation, regression, and mortality among internal sources of invalidity. Except for long-term fatigue effects on researchers as observers (as instrumentation), internally invalidating effects appear to be well-controlled. In general, we have been able to obtain consent to employ the most powerful assignment procedures to generate equivalent comparison groups.[4]

We recognized that there would be "reactive leakage" between the chosen and unchosen children in the schools that we asked to nominate lists of eligible children. The students nominated for the program are from similar ability-track levels, share some classes together, and thus may know each other within the school and neighborhood setting. For our

---

3. See Appendix for a detailed statement of selection procedures given to the schools involved.

4. See the Appendix for a discussion of the technical aspects of obtaining cooperative consent for extensive talent search for generating experimental and control groups, and fair, random assignment procedures in field experiments.

Cleveland sample, therefore, we developed a "nonequivalent control group" of post-matched students from schools post-matched with our chosen schools. Campbell and Stanley label this design as Design #10, The Non-Equivalent Control Group. (This "patching" would not have been necessary had we been able to match impacted schools before pre-matching students within and across schools.) The first control group (C-1) provides data for the most powerful comparison, because we were able to assign students to the experimental and control conditions at random. The second control group (C-2) insured that whatever was done to help the experimental group (Ex)--and which might somehow seep into the school program where both the experimental and control children were enrolled--would not have any confounding effect, because different buildings, teachers, and classes would be involved. This problem is found particularly at the junior high school level; thereafter, students are scattered through various senior high schools.[5]

Chart I outlines the typical cycle of nomination, selection, participation, and evaluation of the participation (or non-participation) in all three cycles of students under study. Within a given cycle, Cohort A includes both the chosen (Ex) and unchosen students (C-1) who were assigned from the chosen school invited to nominate children for the program. Cohort B includes the post-matched students in the post-matched

---

[5] We have now available in the literature ready means for systematic evaluation of the power of alternative designs in controlling plausible rival hypotheses to the research hypothesis. The terminology, labeling, and numbering of alternative designs and the listing of the sources of invalidity are fully developed by Donald T. Campbell and Julian L. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," in Nathan L. Gage, ed., HANDBOOK OF RESEARCH IN TEACHING, Rand MacNally and Co., 1963, pp. 171-246. Campbell and Stanley's discussion also provides to the problem of identifying organizational and extra-organizational constraints and encouragements to given arrangements and design procedures. See also Donald T. Campbell, "Reforms as Experiments," AMERICAN PSYCHOLOGIST, April, 1969, pp. 409-429.

Chart I

A PATCHY INSTITUTIONAL CYCLE DESIGN (MODIFIED DESIGN #15)

Experimental and Control Groups within Cycle (repeat for 3 cycles)		Experimental Treatment (multiple treatments extended over five years)	Measurement	
			Non-Reactive Measures	Reactive Measures
Cohort A (students nominated, & assigned from chosen school) pre-matched*			Design #6: After-Only Fisher Design expanded into Design #15	
Ex Group	Selected randomly to participate	Receives treatment	↑	↑
C-1 Group	Selected randomly to serve as controls	Does not receive treatment	Archival measures from school records	Interviews and paper and pencil measures
Cohort B (students post-matched to Group 1 students)		↓	Design #10: Non-reactive, post-matched, nonequivalent control group expanded into Design #15	
C-1-b Group	Post-matched within chosen school	↓	↓	↓
C-2 Group	Post-matched from unchosen school			

*The first cycle of students in 1964 were assigned by simple random assignment rather than by pre-matching before assignment. Post-matching of nonequivalent matches within the nominated, eligible group has also been made.

schools (C-2) as well as a set of post-matched students in the chosen schools (C-1-b). The latter were sought because no student who was nominated provided an adequate match in the pre-matching effort. Both cohorts are followed over time through archival records of academic performance and by interviews and paper-and-pencil inquiries of the heads of households of the children in the study. The archival checks are less reactive; the other checks may be quite reactive in influencing study results.[6]

The method of pre-matching before random assignment limits our ability to generalize to the total nominated and eligible list of students. Matches cannot be found for all the students who are nominated. We think our initial obligation is to establish the size and direction of differences within the study itself. Later, as programs of talent search become better able to scan the range of students more effectively, generalizable studies could be undertaken on a larger scale.[7]

The Quality of the Matching Process. The best-laid plans for systematic talent search and selection on the basis of pre-matching along selected variables face limiting field conditions and circumstances of one kind or another. We can describe a scale of ability to control the matching process, ranging from matches in which the investigators have full control to those matches which involve post-matching with all its inherent weaknesses. Although a majority of our matches fall in

---

6. See the Appendix for list of instruments used.

7. See the discussion in Donald T. Campbell, "How Regression Artifacts in Quasi-Experimental Evaluations can Mistakenly Make Compensatory Education Look Harmful," ms. J. Hellmuth (Ed.) COMPENSATORY EDUCATION: A NATIONAL DEBATE, Vol. III of THE DISADVANTAGED CHILD (New York: Bruner/Mazel, 1970), pp. 9-11.

category 1 (given below), others are scattered throughout the range. In our interpretation, we shall take account of the quality of the match.

Table I

Types of Matching

0. Not applicable (applies to nominated subjects who weren't matched).

1. Paired by selection team (Yinger, Ikeda, Laycock, Antes) from original nominated list; randomly chosen from pairs, one as experimental student, one as control 1 student.

2. Experimental group chosen by team randomly from entire nominated list; control 1 student matched by team later from original nominated list.

3. Experimental student chosen by team, but not randomly from entire list; control 1 student matched by team later from original nominated list.

4. Experimental group chosen by school authorities; control 1 student matched by team later from original nominated list.

5. Experimental student chosen by team, but not randomly from entire list; no satisfactory match from original nominated list; control 1 student matched by team later from further suggestions or nominations from school authorities.

6. Experimental student chosen by team, but not randomly from entire nominated list; no satisfactory match from original nominated list; control 1 student matched by team later from school records of students not originally or later nominated or suggested by school authorities.

7. Control 2.

OPERATIONAL MEASURES OF INITIAL AND LONG-TERM PROGRAM IMPACT

We made a systematic effort to measure the impact of our program by development of instruments which had some chance of uniform application to both chosen and unchosen nominees. Moreover, these procedures had the potential of being non-obtrusive and non-reactive in evaluating program impact. The following instruments were developed and employed.

1. The Pupil Data Form. This form was applied uniformly to both chosen and unchosen children to obtain the total list of nominated students from the schools. Pre-matching before random assignment was completed from these forms, with criteria for selection into the eligible pool of students being determined on the basis of selection criteria reported above.

2. The Student's Permanent Record Form. After a careful review of the permanent record forms of students in the five school districts and any variant forms from other schools to which the students transferred, we developed a uniform record form to retrieve selected academic information. Pre-program measures and post-program measures are obtained from this record on the following variables:

a. Grades over the study period, usually semester by semester (on a 4-point scale).

b. Test scores on achievement and intellective functioning (usually converted to stanine scores).

c. Academic track level, where applicable, and related information on quality and direction of schooling (college preparatory in emphasis or not).

Our rate of completion in obtaining parallel information on grades, tests, and quality of schooling is 441 (93%) out of 474 cases. For various subgroups, the rate of return was as follows:

Table 2

ARCHIVAL INFORMATION

	1964			1965			1966			Total		
	Data	n	%	Data	n	%	Data	n	%	Data	n	%
Ex	49	56	88	66	70	94	63	69	91	178	195	91
C-1	52	56	93	65	70	93	62	69	90	179	195	92
C-2	23	24	96	26	33	79	35	37	95	84	94	89

We believe that this is an excellent record, considering the field conditions under which such procedures must be carried out.

3. Parents' Interview Schedule. To maintain consistency and to speak with the individual most likely to be involved in the education of the children, we sought the mother or allied woman of the house as the primary respondent. The following variables were primary in interviewing this person:

a. A description of family resources and conditions as they related to ability to support the child financially and in guidance and instruction for academic achievement.

b. The adult's dreams and aspirations for the child in educational and occupational career matters.

c. The adult's access and sense of access to reference persons and groups who could support, guide, and instruct the child involved to achieve well in educational and occupational terms.

d. Motivational-attitudinal factors of achievement and of optimism-pessimism in keeping at the task of educational and occupational achievement.

e. Testimonial data on the impact of the S.O.P. effort, as perceived by the adult.

These interviews were obtained through a 26-month period (June, 1968 through July, 1970).

Table 3

INTERVIEWS

	1964			1965			1966			Total		
	Completed	n	%	Completed	n	%	Completed	n	%	Completed	n	%
Ex	49	56	88	64	70	91	62	69	90	175	195	90
C-1	37	56	66	50	70	71	52	69	75	139	195	71
C-2	18	24	75	30	33	91	33	37	89	81	94	86
Refused[8]												
Ex	3			0			1			4		2
C-1	1			7			3			11		6
C-2	4			2			2			8		9
No Contact (moved out of area, or not locatable through any public source)												
Ex	4			6			6			16		8
C-1	18			13			14			45		23
C-2	2			1			2			5		5

4. The Mailed Questionnaire. Our most recent information comes from a mailed questionnaire, sent to both participant and non-participant households. These data were secured through four mailings. Where necessary, we then supplemented this source with information from school counselors, who had obtained parallel information from their annual interviews with students on matters of academic and occupational planning.

---

[8] Of the Refusals, 15 were White, 8 were Black, 0 were Americans of Spanish-speaking descent.

The following statistical breakdown is reported on the rate of return by mailing and by substitute filing of reports by counselors and other archival information:

Table 4

QUESTIONNAIRE AND COUNSELOR RETURNS

	1964		1965		1966		Total	
Ex	(56)	%	(70)	%	(69)	%	(195)	%
Questionnaire	45	80	59	84	56	81	160	82
Counselors and other sources	11	20	11	16	11	16	33	17
No info	0		0		2	3	2	1
C-1	(56)		(70)		(69)		(195)	
Questionnaire	33	59	48	69	44	64	125	64
Counselors and other sources	22	39	18	26	20	29	60	31
No info	1	2	4	6	5	7	10	5
C-2	(24)		(33)		(37)		(94)	
Questionnaire	17	71	27	82	23	62	67	71
Counselors and other sources	6	25	5	15	13	35	24	26
No info	1	4	1	3	1	3	3	3

Given the backgrounds of the households, we believe that this rate of return is remarkable.[9]

---

[9] It probably matches the rate in Sewell's study cited above, if the returnees are equated by socio-economic status.

SUMMARY OF METHODOLOGICAL PROCEDURES

Further comments on method will be made in conjunction with our interpretations of the data of our study. In broadest outline, our research design calls for a comparative study of educational attainment of experimental and control groups, matched as closely as possible. We shall take account of the quality of the match in the interpretations. In addition to study of the impact of the summer program and follow-up activities, we shall examine the effect of other intervening variables (networks of significant others, most importantly), using them for specification or other forms of elaboration, in the Lazarsfeld sense of those terms. Then, by regression analysis and other forms of statistical treatment, we shall attempt to indicate the sources of any observed differences in the dependent variable. At some points, the dependent variable will be stated as a single index number; at other points we shall use specific dependent measures--grades, test scores, and number of years of education, for example.

In schematic terms, then, our design is as follows:

		Input Variables (tendencies and supportive opportunities of observed children)	Intervening Variables		Outcome Variables (levels of educational attainment)
			Program and follow-up	Other Intervening Variables	
Matched	Experimental Children	X	X	X	X
	Control Children	X	0	X	X

Chapter Three: BACKGROUNDS OF THE MIDDLE STARTERS

Who were the children in the program? What kinds of resources did they bring to their educational experience? How much family support could they count on? Before undertaking the analysis and interpretation of our empirical findings (in Chapters 4 and 5), we shall describe briefly the background characteristics of the students in the Special Opportunities Program.

It was our aim to work with students who showed some promise of academic success, as indicated by the usual measures, but who also carried some burdens. We did not believe that our experimental stimulus would be strong enough or persist long enough to overcome extreme academic deprivation. At the same time, if we selected only those who were already ahead of their peers (the kind of "good students" whom teachers would have rewarded if we had asked them to nominate only their best), we could scarcely test our ability to modify academic performance. We wanted to see if careful attention both to individual skills, motives, and aspirations, and also to socio-cultural supports from schools, families, and the college could significantly alter the expected sequence of events. We sought, therefore, to select students with some potential, but with academic handicaps sufficiently large that without special "intervention" most would not go beyond high school.

The following data may give a picture of the range of factors influencing the experimental group as they entered the program.

THE SIZE AND GENERAL COMPOSITION OF THE EXPERIMENTAL GROUP OF PUPILS

First, who were the pupils who were picked for a "special opportunity"? Table 1 displays their numbers, arranged by the year in which they entered the program, their sex, and their race.

Table 1: Distribution of Pupils
by Year of Participation, Sex, and Race

Trait	Frequency by Year			
	1964	1965	1966	Years combined
Sex				
male	38	45	46	129
female	18	25	23	66
Total	56	70	69	195
Race				
White	14	17	15	46
Spanish-Mexican	1	3	-	4
Puerto Rican	3	3	4	10
Negro	37	46	48	131
Mixed	1	1	1	3
Amer. Indian	-	-	1	1

Thus the total group of 195 pupils were two-thirds male and two-thirds Negro. These proportions reflect a deliberate policy decision, to seek out the pupils who would have to bear the greater burden of occupational responsibility during adulthood (the boys) and the group that was most numerous among the severely disadvantaged (the Negroes).

Another policy decision dictated the choices shown in Table 2. Oberlin arranged to cooperate with four school districts close to the campus, and with the Banneker district of the St. Louis schools.

Table 2: Cities from Which Pupils Came

City	Frequency by Year			
	1964	1965	1966	Combined
Cleveland	24	33	37	94
Elyria	6	7	5	18
Lorain	9	10	10	29
Oberlin	5	8	5	18
St. Louis	12	12	12	36

At the time pupils first came to Oberlin they were between the seventh and eighth grades (although there were a few pupils in the first group, 1964, who were between the eighth and ninth grades). Table 3 gives a distribution of ages, showing a range from about twelve and one-half years to about fourteen and one-half years, around an average of thirteen years. The pupils were on schedule, as one would expect from the typical lock-step progress through the grades in public schools.

Table 3: Distribution of Ages at Time of Entering Program

Age (in months)	Frequency by Year			
	1964	1965	1966	Combined
184-5	-	1	-	1
182-	-	-	-	-
180-	-	1	-	1
178-	-	-	-	-
176-	1	2	1	4
174-	-	1	1	2
172-	3	1	2	6
170-	3	-	1	4
168-	1	1	3	5
166-	8	3	2	13
164-	4	3	4	11
162-	2	6	5	13
160-	4	6	-	10
158-	11	14	21	46
156-	9	9	10	28
154-	7	12	6	25
152-	2	10	12	24
150-	1	-	-	1
148-9	-	1	-	1

BASES FOR CHOOSING EXPERIMENTAL PUPILS

From among a large pool of disadvantaged pupils, the aim was to find those who had sufficient potential to overcome their disadvantage if given help. The cooperating schools nominated pupils whose tested aptitude and achievement were superior, and whose other traits (motivation, emotional and social maturity) would reinforce talent and respond to assistance. For these latter traits Oberlin depended upon teacher recommendation and similar testimonials. For aptitude and achievement data, scores were provided from the regular batteries routinely administered in all the cooperating school districts throughout a pupil's career.

Table 4 shows the pattern of tested intelligence. Many specific tests were used, in varying combinations in the several districts (CTMM, WISC, Kuhlmann-Anderson, Pintner Stanford-Binet, Lorge-Thorndike, Otis, Terman-McNemar, and the Cleveland test of Probable Learning Rate were the major ones). Since scores were not directly interchangeable, they were converted to stanines, for all scores except those from St. Louis. Normative data were insufficient for a transformation of the St. Louis scores. These scores, in their original I. Q. form, are in Table 4a.

Table 4: Tested Intelligence Level at Time of Entry

Intelligence score (stanines)	Frequency by Year			
	1964	1965	1966	Combined
9	8	3	2	13
8	8	6	5	19
7	6	9	10	25
6	7	9	8	24
5	4	8	10	22
4	1	2	6	9
3	-	2	1	3
2	-	-	-	-
1	-	-	-	-
Mean	7.2	6.3	6.0	6.5
Standard deviation	1.4	1.3	1.2	1.3

(Note to Table 4)

These stanines do not include data from St. Louis. See Table 4a.

Table 4a: Intelligence Test Scores (St. Louis only)

Score	Year			
(I.Q.)	1964	1965	1966	Combined
125-29	-	1	-	1
120-	4	-	1	5
115-	1	2	1	4
110-	3	4	-	7
105-	1	3	2	6
100-	2	1	2	5
95-	1	-	3	4
90-	-	-	3	3
85-89	-	1	-	1
Total	12	12	12	36

These scores in Tables 4 and 4a show that most pupils who entered were superior in tested intelligence. This superiority is important to note, in the midst of contemporary arguments that standard tests of intelligence systematically under-measure pupils from lower-class backgrounds. These pupils were recommended because of such backgrounds; they showed nonetheless a strong potential even against allegedly alien norms. It should also be noted that the 1965 and 1966 groups had lower average scores--an indication of our interest in testing the effects of the program on individuals of lesser "ability," as measured by these tests.

Similarly, Table 5 and Table 6, showing average scores on achievement batteries and average class marks, validate the policy aim to find students who were doing satisfactory or superior work.

Table 5: Mean Grade-Level Scores on Achievement Tests at Time of Entry

	1964	1965	1966
Mean Grade-Level Score	8.6	8.0	7.6

Table 6: Mean Class Marks at Time of Entry

School Grades	Frequency of Mean Class Marks (A=4)			
	1964	1965	1966	Combined
7th	2.8	2.7	2.7	2.7

The typical pupil, then, had better than average academic aptitude, was making satisfactory progress at grade or above on standard achievement tests, and had marks averaging B-, when he entered the program. About one-fifth, however, were below average, by the national or local norms being used, and another one-fifth of the participants were below average in one or another of the scales.

HOME BACKGROUND

Although the home was not analyzed as part of the basis for choosing the experimental pupils, it is important to find out what sort of background the pupils came from. Table 7 shows that the average pupil had a much larger than average number of siblings:

Table 7: Siblings in the Home

Number of Siblings	Year			
	1964	1965	1966	Combined
17	-	1	-	1
13	-	1	-	1
11	2	-	2	4
10	1	3	1	5
9	2	2	2	6
8	1	2	1	4
7	3	5	3	11
6	4	3	13	20
5	3	8	10	21
4	10	12	9	31
3	10	15	8	33
2	12	10	7	29
1	7	7	10	24
0	1	1	3	5
Mean	3.9	4.2	4.2	4.1

Since the Oberlin program could provide only a relatively small amount of help, even when combined with what a pupil's regular school could offer, it is necessary to find out about the parents. Tables 8 and 9 show a significant pattern: the homes were managed by the mother to a far greater extent than is typical in other social and ethnic strata.

Table 8: Father's Status in the Home

Father's Status	Frequency by Year			
	1964	1965	1966	Combined
Real father in home	34	44	37	115
Stepfather in home	5	6	3	14
Other adult male in home	1	1	4	6
Father dead; no other adult male in home	-	2	2	4
Father not in home, but alive; no other adult male in home	15	17	20	52
Father unknown; no other adult male in home	1	0	3	4

Table 9: Mother's Status in the Home

Mother's Status	1964	1965	1966	Combined
Real mother in home	50	66	62	178
Stepmother in home	0	2	0	2
Other adult female in home	1	2	7	10
Mother dead; no other adult female in home	2	0	0	2
Mother not in home, but alive	3	0	0	3

Table 10 gives the outline of parental education, showing important contrasts between father and mother.

Table 10: Education of Father and Mother

Highest Level of Schooling	Frequency by Year							
	1964		1965		1966		Combined	
	Father	Mother	Father	Mother	Father	Mother	Father	Mother
High school graduate or higher	15	23	14	26	19	24	48	73
9-11 years	9	18	10	22	11	24	30	64
5-8 years	9	6	15	11	8	8	32	25
0-4 years	2	1	1	1	2	2	5	4
Not applicable; or information lacking	21	8	30	10	29	11	80	29

Directly related to the disadvantaged background that was the prime desideratum, the parents' occupation is a crucial fact. Tables 11 and 12 show the distribution into standard classification categories:

Table 11: Father's Occupational Skill Level

Occupational Skill Level	Frequency by Year			
	1964	1965	1966	Combined
Sales, clerical managerial, professional	2	2	5	9
Skilled worker, foreman	8	10	9	27
Semi-skilled, service, household worker	15	21	29	55
Laborer, farm laborer	13	22	8	43
Not applicable; or information not available	18	15	18	51

Table 12: Mother's Occupational Skill Level

Occupational Skill Level	Frequency by Year			
	1964	1965	1966	Combined
Sales, clerical, managerial, professional	5	5	6	16
Skilled worker, foreman	1	2	0	3
Housewife, laborer, service worker	44	61	58	163
Not applicable or information lacking	6	2	5	13

Related to occupational skill, and particularly important because so often the mother was both wage-earner and home-manager, is her status in these two situations. Table 13 shows the distribution of mothers according to how much responsibility they took outside the home.

Table 13: Mother's Occupational Status

Occupational Status	Frequency by Year			
	1964	1965	1966	Combined
Full-time employment	9	18	22	49
Part-time employment	2	0	2	4
Housewife	37	50	41	128
Not applicable; or information lacking	8	2	4	3

Finally, since churches provide considerable support to the aspirations of many families, the data in Table 14 show that most parents of the experimental pupils belong to a church, and that it is predominantly a Protestant church.

Table 14: Parents' Religious Preference

Religious Preference	Frequency by Year								
	1964		1965		1966		Combined		
	Father	Mother	Father	Mother	Father	Mother	Father	Mother	Total
Protestant	26	35	36	55	35	53	97	143	240
Roman Catholic	14	10	4	6	5	8	23	24	47
Jewish	-	1	2	-	-	-	2	1	3
No information	16	10	28	9	29	8	73	27	100

Chapter Four: TESTING AN HYPOTHESIS BY USE OF MATCHED PAIRS

By the research reported here we are attempting to test the following hypothesis:

> Among culturally deprived youngsters who have finished the seventh grade, an intensive but pleasant summer's experience on a college campus, followed by long-term counselling, newsletters, reunion periods, supplemental academic training, conferences with parents, and other "follow-up" stimuli through the next several years, will significantly affect their academic performance, the likelihood of their staying in school, their occupational plans, and the rate of college attendance.

The research design by which we are testing this hypothesis is less commonly employed than designs appropriate to the small group laboratory, with its possibilities for rather elegant controls, or the survey, with its possibilities for careful sampling even of national populations. Our design, however, shares some aspects with both of the others. Like the small-group laboratory research, ours is a true experiment; and like the survey, it is "in the field," a study of "real life" conditions.[1] This does not mean that we have the best of both worlds, for a field study that attempts a true experimental design faces difficulties of its own, which we shall examine. It does have the great advantage, however, of both relevance and an experimental structure.

To some degree, approaches to the study of human behavior are reflections of what Robin Williams has called "types of scientific conscience."

> To the historical and cultural conscience it is above all important that the object of study be historically and culturally important.... a conscience of this type would insist

---

[1] For a discussion of the problem of realism in sociological research see Thomas Drabek and Eugene Haas, "Realism in Laboratory Simulation: Myth or Method?" Social Forces, March, 1967, pp. 337-346.

2

> upon intimate familiarity with a wide range of materials, and place a high value on erudition.
>
> To the "clinical" sociologist, on the other hand, a primary virtue is detailed and sensitive fidelity to the comples, immediate situation. His anxiety dreams are likely to be studded with horrid fancies of having "torn a fact out of context" or, perhaps worse, having "generalized beyond his data." His conscience is clear and his disposition sunny when after a long experience of immersion in a factory work group or a boy's gang he completes a vivid naturalistic description of complex behavior and its complex motivation. In his harsher moments, he may describe the historical or cultural sociologist as an "arm-chair theoriest," the experiment as "artificial," and the survey as "crude" and "mechanical."
>
> To persons in the logico-experimental group, the ideal study is the highly controlled experiment or the sample survey, complete with scales, scores, probability samples, and possibly electronic computers.[2]

Our "scientific consciences" have led us closest to the last type mentioned by Williams. By the use of carefully matched experimental and control groups we shall test a causal hypothesis. At the same time, we hope to maintain something of the "clinical" style, by locating our facts in the context of school and community, and of the "historical" style, by recognizing that the events we are studying are part of a major transformation of American society and cannot be understood without awareness of their relationship to this transformation.

Because we have undertaken a six-year research program in the field, we are confronted with experimental problems that are much less likely to disturb short-time laboratory research. These are the prices we pay for realism. There has been some loss of respondents. At various points we shall note that information on one or both members of a matched pair is lacking, because at some point in the six-year period they refused to answer further inquiries or moved to addresses that we have been unable to learn about. Although this loss represents a small proportion of the original group, it raises questions of selective loss with which we shall have to deal. There is some problem of "leakage," since our experimental and control children attend the same schools. Although it seems unlikely that more than a tiny fraction of the stimulus effect of the summer program and follow-up activities can be transmitted to the control children, through their contact with experimental children, we shall attempt to measure that fraction by reference to a second control group in one of our school systems. This second control group is made up of children who attend schools similar to but not identical with those attended by children who participated in our program. This second group of controls has also been individually matched with experimentals.

---

[2] Robin M. Williams, Jr., "Continuity and Change in Sociological Study," AMERICAN SOCIOLOGICAL REVIEW, Dec., 1958, pp. 622-623. Today he would scarcely say "possibly" with reference to computers.

A longitudinal study faces the problem of "history." How much are observed changes the products of the program, how much the results of general developments in society or in the communities studied? This problem is increased if historical forces are having general effects somewhat similar to the effects of the program, which is true in our case. Ideally, comparison of the experimental and control groups should permit us to separate historical from experimental influences. This is possible, however, only if control and experimental pairs have been carefully matched. Almost every problem in fact, in a research design that employs a field study is reduced if matching has been done carefully. Precise matching, however, is one of the most difficult tasks in sociological research.[3] Many variables may affect the influence of the experimental stimulus, and must therefore be matched for; theory may be insufficiently developed to guide one to the critical variables or to their appropriate weighting; measurement of some of the variables is often rough; interaction effects are highly complex and not well understood; the pool from which both experimental and control groups must be drawn is often quite small, making precise matching difficult; and randomization may be unavailable, or only partially available.

---

[3] The next several pages are based on our paper "Treating Matching as a Variable in a Sociological Experiment," AMERICAN SOCIOLOGICAL REVIEW, Oct., 1967, pp. 801-812.

Despite the fact that such problems are common in sociological studies, even when experimental designs are used, there is relatively little discussion of ways to compensate for departures from ideal matching conditions.[4] In the laboratory, a judicious combination of matching and randomization can often approximate the desired pattern. Various restraints on field research, however, may lead to wider deviation from the model, leaving the interpretation of the results in doubt.

---

[4] The discussion of matching and randomization in Ernest Greenwood, EXPERIMENTAL SOCIOLOGY, A STUDY IN METHOD, New York: Crown Press, 1945, is still highly useful, as is the treatment by F. Stuart Chapin in EXPERIMENTAL DESIGN IN SOCIOLOGICAL RESEARCH, rev. ed., New York: Harper and Bros., 1955. A classic statement of the use of randomization is found in R. A. Fisher, THE DESIGN OF EXPERIMENTS, 2nd ed., Edinburgh: Oliver and Boyd, 1937. Several recent texts have brief treatments of matching. See, for example, A. L. Edwards, EXPERIMENTAL DESIGN IN PSYCHOLOGICAL RESEARCH, New York: Rinehart and Co., 1950, chap. 14; Matilda W. Riley, SOCIOLOGICAL RESEARCH, New York: Harcourt, Brace, and World, Inc., 1963, pp. 614-620, 635-637; Claire Selltiz, Marie Jahoda, Morton Deutsch, and Stuart W. Cook, RESEARCH METHODS IN SOCIAL RELATIONS, rev. ed., New York: Henry Holt and Co., 1959, pp. 98-99, 102-108, and 138-139. One of the best empirical works, using matching, is the survey study of Ronald Freedman and Amos Hawley, "Unemployment and Migration in the Depression (1930-1935)," JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION, 44 (June, 1949), pp. 260-272. For a recent matching study, see Ralph Schwitzgebel, STREETCORNER RESEARCH, Cambridge: Harvard University Press, 1964.

THE ROLE OF MATCHING IN SOCIOLOGICAL EXPERIMENTS

To impute causal effects to experimental treatments, one seeks to establish equivalence in initial characteristics between the experimental and control groups, which are arranged for systematic comparisons. There are several possible methods for obtaining this initial equivalence. The most general procedure, once thought to produce rather precisely equivalent comparison groups, is pre-testing, but it is now viewed as less than adequate because pre-testing may sensitize respondents to the experiment, thus modifying its influence. Unless one wants to measure the effects of pre-testing itself, the recommended procedure now is to assign experimental and control groups randomly from a given population pool--the Fisher "after-only" experimental design. If a large population is available, this design is to be preferred over other arrangements. When the number of eligible participants is small, however, and the population likely to be heterogeneous, random assignment without initial controls on the likely heterogeneity can lead to the appearance of no difference between control and experimental subjects, or it can lead to the appearance of a spurious difference. Such results are due to interactions between the experimental stimulus and initial, uncontrolled differences among respondents.

There are various ways to avoid or reduce these errors. If there are fairly clear-cut strata in the population, its division into more nearly homogeneous sub-sets is a valuable procedure. Where the available nominations for assignment into control and experimental groups is small, however, such stratification procedures are not feasible. The alternative is to match pairs of subjects before random assignment. Although the matching process reduces the extent to which findings can be generalized to a wider population, because of the loss of non-matchable cases, it can help to insure initial equivalence.

In this context, we should note that the use of matching in conjunction with random assignment to experimental and control groups increases the precision of one's interpretation.[5] Even with this increase in precision, however, problems of analysis and interpretation remain. The task is to choose a procedure that will take account of the multiple effects of the matched variables as they interact with the experimental treatment to produce given results.

To know whether our program of educational intervention has produced the desired effects, we need measures, through several years, of the educational performance not only of the children in the program but also of children in the control groups. Most of the 390 students who make up the core of the study had just finished the seventh grade at the time we began our observations.[6] They come from five cities and fourteen schools. To determine whether the summer and follow-up program have promoted educational aspirations and performance, we asked the several school systems involved to nominate at least twice as many children as they were entitled to send to the program. We paired the children on the basis of several criteria, indicated below, and then randomly assigned one of each pair to the experimental group and one to the control group. Such random assignment of matched pairs is a form of selection designed to heighten one's confidence in interpreting the results of the experimental variable.

---

[5] Campbell and Stanley suggest that ". . . matching can be recognized as a useful adjunct to randomization but not as a substitute for it; in terms of scores on the pretest or on related variables, the total population available for experimental purposes can be organized into carefully matched pairs of subjects; members of these pairs can then be assigned at random to the experimental or to the control conditions. Such matching plus subsequent randomization usually produces an experimental design with greater precision than would randomization alone." Donald T. Campbell and Jullian L. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," in Nathan L. Gage, ed., HANDBOOK OF RESEARCH IN TEACHING, Chicago: Rand MacNally and Co., 1963, p. 219.

[6] The number of matched pairs will range downward from 195, depending upon the availability of certain kinds of information. We have data on additional students from the same schools, however. Thus N for some comparisons will be over 700.

Several departures from the ideal matching situation were, however, inevitable. Such information as father's occupation, or a full record of school grades, sometimes was lacking. In a few instances, the school personnel, either through misunderstanding, or as a result of a powerful (and understandable) urge to reward their "best" children, sent us a list of appointed children and of "alternates"--a procedure that restricted the use of randomization. Most important of all, the lists from which we drew the selections were seldom more than twice as large as the number to be selected, limiting the range of variables we could take into account.[7]

Despite these problems, which are not uncommon in field research, we undertook to approximate individual precision matching as closely as we could. Ideally, we would have matched, for example, a 13-year-old

---

[7] It should be noted that from the point of view of external validity, there is some advantage in having to match nearly 100 percent of the individuals. There can be very little exclusion of those who are difficult to match--an exclusion that can produce a systematic bias. (See Ronald Freeman, "Incomplete Matching in Ex Post Facto Studies," AMERICAN JOURNAL OF SOCIOLOGY, [March, 1950], pp. 485-587.) But there is also the disadvantage that the match will necessarily be rougher when all subjects must be included. The question is, how much rougher, and how is the imprecision to be measured? For a discussion of questions of external and internal validity, see Donald T. Campbell and Julian C. Stanley, _op_. _cit_., pp. 171-246.

Negro boy from a given school in St. Louis with another 13-year-old Negro boy from the same school, both of them having been recorded as having C+ academic averages, IQs of 108, similarity on other tests, two older brothers and three younger sisters, fathers who were semi-skilled workers living in the household, mothers who were housewives, each of whom had finished the ninth grade, etc., etc.

In practice, we matched for ethnic group, sex, class in school, city, father's place in the household, roughly for academic and intelligence measures--and then randomized. This is the point at which the matching process usually stops. Researchers and their reviewers then tend to take one of two positions: either they underline the number and seriousness of the compromises from the ideal model, note the need for great caution in interpreting the data, and call for further research; or they note the great care used in matching as closely as possible, emphasize the beneficent statistical effects of randomization--and hope for the best.

The Quality of the Matching Process--An Index of Congruence. Not completely comfortable with either of these choices, we have designed an index of congruence which can tell us how closely our operations produced matched pairs of individuals. Since this index is a rough indicator of equivalence, and since the variables noted below are likely to represent highly inter-correlated influences, producing both interaction and main effects of unknown magnitude, this index will be pitted against alternative modes of analysis to determine its utility in identifying effects associated with both the matched variables and the experimental treatment, compared with the more standard alternate procedures.

There are several arbitrary qualities to an index of congruence among matched persons or units. Which variables should be employed in the match (to control unwanted influences), and which should be allowed to vary? If the index is to be employed as a scale of initial equivalence or non-equivalence, how should each variable be weighted? How large should the categories be, in the variables that are quantified, to indicate similarity or difference? How should interaction effects be treated and taken into account?[8] This last is a particularly troublesome question, because the empirical support, and the theory to guide choice of important interaction effects, are not well developed. It seems highly plausible, for example, to hold that the influence of the variable "mother works outside the home" or "does not work outside the home" is conditioned by another variable--"father present" or "father absent." Its effects also may be modified by presence or absence of older siblings in the home, by ethnic groups, and by other variables.

We shall attempt in various ways to take such interaction effects into account and to remain aware of the arbitrary aspects of our index. With these cautions in mind, we describe here the eighteen variables selected for the index, their weights, and the operations by which we assigned the weights:

---

[8] See Hubert M. Blalock, Jr., "Theory Building and the Statistical Concept of Interaction," AMERICAN SOCIOLOGICAL REVIEW, 30 (June, 1965), pp. 374-380.

INDEX OF CONGRUENCE

		Range of Possible Point Difference Between Matched Pairs
1. Ethnic or Racial Group		0 to 4 points
White-White Negro-Negro "Spanish"-"Spanish" Etc.	= 0	
"Spanish"-White Negro-"Spanish" Other nonwhite-White Other nonwhite-Negro	= 2	
Negro-White	= 4	
2. Age		0 to 2 points
Under 6 months difference	= 0	
6 months to 11 months difference	= 1	
Over 12 months difference	= 2	
3. Sex		0 to 3 points
Same	= 0	
Different	= 3	
4. City of residence		0 to 2 points
Same	= 0	
Different	= 2	
5. School attended		0 to 1 point
Same	= 0	
Different	= 1	
6. Status of father in home[9]		0 to 5 points
(Points are calculated by finding the numerical difference between		

---

[9] The effect of calculations based on variables 6-8 is to produce an "advantage score" relative to adult male models. The sharpest contrast possible would be between a child whose father is unknown and one who lives with his real father, who is a white-collar worker or semi-professional with at least a high school education.

members of the pair, using
the following scale. Use
highest number applicable.)

0. Father unknown. No other adult male in home
1. Father not in home, but alive; no other adult male in home
2. Father dead; no other adult male in home
3. Other adult male in home
4. Stepfather in home
5. Real father in home

7. Father's occupation (or that of other adult male) 0 to 4 points

(Calculate as above)

0. No adult male in home
1. Unskilled
2. Semiskilled
3. Skilled
4. White-collar or semi-professional

8. Education of father (or that of other adult male) 0 to 4 points

(calculate as in item 6)

0. No adult male in home
1. Father or other adult male has less than five years of education
2. 5 to 8 years of education
3. 9 to 11 years of education
4. High school graduate or higher

9. Status of mother in home 0 to 4 points

(Calculate as in item 6)

0. Mother not in home, but alive
1. Mother dead; no other adult female in home
2. Other adult female in home
3. Stepmother in home
4. Real mother in home

10. Occupational status of mother (or that of other adult female) during most of preceding year — 0 to 3 points

(Calculate as in item 6)

0. No adult female in home
1. Housewife
2. Part-time work
3. Full-time work

11. Occupational skill of mother (or that of other adult female) — 0 to 3 points

(Calculate as in item 6)

0. No adult female in home
1. Housewife, unskilled, or semi-skilled
2. Skilled
3. White-collar or semi-professional

12. Education of mother (or that of other adult female) — 0 to 4 points

(Calculate as in item 6)

0. No adult female in home
1. Mother or other adult female has less than five years of education
2. 5 to 8 years of education
3. 9 to 11 years of education
4. High school graduate or higher

13. Sibling patterns — 0 to 6 points

(Several of these statements may apply to a pair. They are additive, up to six points)

a. Only child in both cases = 0
   One an only child, the other not = 1
b. Both eldest children = 0
   One the eldest, the other not = 1
c. Both have or neither has siblings of same sex = 0
   One has siblings of same sex, the other does not = 1
d. Both have or neither has siblings of opposite sex = 0
   One has siblings of opposite sex, the other does not = 1

e. Difference of one or no difference in number of siblings = 0
2 or 3 difference in number of siblings = 1
4 or more difference in number of siblings = 2

14. Intelligence Quotients 0 to 8 points

(Different tests and dates having been equated, where necessary, by translation into stanines)[10]

Same stanine = 0
Different stanine = 1 to 8

15. Achievement tests 0 to 8 points

(Different tests and dates having been equated where necessary by translation into stanines)

Same stanine = 0
Different stanine = 1 to 8

16. Grades (Average of 7th grade scores in "solid" subjects.

4=A; 3=B; 2=C; 1=D) 0 to 3 points

0- .49 difference = 0
.50- .99 difference = 1
1.00-1.49 difference = 2
1.50 and more difference = 3

17. Religion of parents, identification 0 to 2 points

Similar background (parents of both Catholic, or both Protestant, e.g.) = 0

---

[10] Among various methods of making comparisons among persons who have taken different tests, use of their stanine scores seems most generally valuable. Most of our pairs took the same test during the same year, so inter-test comparisons are uncommon. Although there is the theoretical possibility of an 8 point difference (if one of a pair fell in stanine 1, the other in stanine 9), the empirical difference seldom exceeds 3.

Mixed background (match between one of parents from each set: e.g., both mothers Catholic, but one father Protestant, one Catholic)	= 1
Different religious back-ground	= 2

18. Religion of parents, importance of 0 to 2 points

(Calculate as in 6)

1. Religion strongly emphasized in home
2. Some emphasis on religion
3. Little emphasis on religion

The effect of these operations is to produce a scale ranging from 0 to 68. A score of 0 indicates a "perfect" match within the definitions of the variables used; 68 indicates the maximum possible measured difference between two persons in the population with which we are dealing. These are arbitrary numbers, of course, which become meaningful only when used comparatively. It is clear that any score near the high end of the range would indicate a serious flaw in the matching process. In an illustrative case, a score of 10 represents the following departures from a "perfect" match: a 7 months' difference in age (1), real father in one home, another adult male (grandfather) in the other (2), one adult male with four years of schooling, the other with nine (2), one mother with eight years of schooling, the other with ten (1), one of the pair without sisters and with 2 fewer siblings (2), one stanine difference in IQ (1), and a .70 difference in average grades (1).[11]

---

[11] At first we used negative numbers for each departure from a good match, to emphasize to ourselves the need for caution in interpreting our results. Positive numbers, however, are probably simpler to use. We need only emphasize that the higher the score, the poorer the match.

In the light of such comparisons, before concluding that the experimental variable has produced any observed results, we shall trace its relationship with the matching index. In statistical terms, if a coefficient of correlation between the matching index score and a score indicating differentials in educational achievement between experimental and control pairs proves to be high, we will not conclude, without support from other procedures, that the experimental stimulus is producing the observed differences. Conversely, the lack of correlation between the matching index and measures of change will support our hypothesis. Thus line A, in Figure 1 below, would cast serious doubts on the proposition that differences in "matched" pairs were produced by the program. Line B, however, would be an indication of the efficacy of the program.[12]

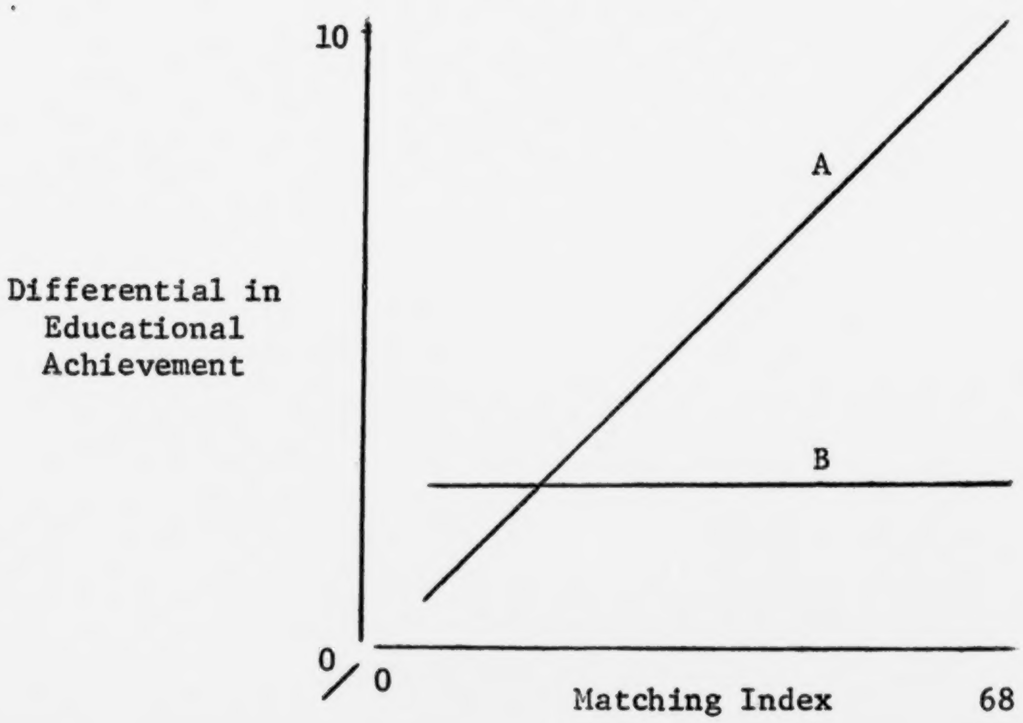

Figure 1.--Illustrative Correlations Between "Matching Index" Scores and "Differential Achievement" Scores, Among Matched Pairs

---

[12] This statement assumes a linear relationship. If, in fact, the relationship is curvilinear, in the form of a U-shaped curve, for example, more complex interpretations will be required.

"IDEAL" VERSUS "DIRECTIONAL" MATCHING

The procedures described so far are based on a very severe definition of a good match. Low scores can be obtained by these methods only if the original pool from which the experimental and control groups were selected was large, thus making it more likely that one can find matched combinations, and only if the experimental and control groups are also large, giving free play to the process of randomization. The procedures imposed by this definition may, in fact, be too stringent for some purposes. The model of a good match that it sets before the researcher does not permit one to take account of the fact that theory may guide one to possible "cancelling out" effects of two mis-matched variables. If one can assign direction to the influence of some variables, he may be justified in saying that a person with a score of -1 on variable A and +1 on variable B is well matched with a person whose scores on the same variables are +1 and -1. In terms of the model we have described above, such a pair would have two negative scores, but using a directional model of matching, one would record this as a zero score.

On the basis of previous research and theoretical considerations, we can indicate quite clearly certain predictors of academic success. We shall consider five: (1) a composite score on father's status (derived from information on his educational level, occupation, and presence or absence from the home); (2) a similar composite score for mother's status; (3) an intelligence test score; (4) various achievement test scores; and (5) average grades in school. To illustrate the procedure, we shall record here only whether the experimental subject is similar to his control partner on a predictor (0), superior to him--that is, has a stronger predisposition to academic success (+1), or inferior to the control subject (-1). In a given case, the "ideal" model for matching might indicate that a pair of subjects was poorly matched on four of the five variables, while a "directional" model would record them as "perfectly" matched:

	Ideal Model		Directional Model	
Variable A	+1		+1	
Variable B	-1		-1	
Variable C	0		0	
Variable D	+1		+1	
Variable E	-1		-1	
	4	(four mismatches)	0	(advantages and disadvantages of predispositions cancel out)

Although the directional model of matching uses existing theory more fully than the ideal model, there is a serious risk involved in using the former: interaction effects are disregarded. We have not taken account directly of interaction effects in preparing the longer matching index, as we have noted, but they are controlled for indirectly in a good match. If two persons are similar on measures of variables A, B, and C taken singly, they are similar in any possible interaction effects among these variables.[13] If they are "similar," however, as a result of algebraic addition of dissimilar series, one may not assume that interaction effects are controlled. It may be well to illustrate this contrast between the ideal and directional models. Assume, for example, that to have a parent who has finished high school is a positive influence on academic achievement of a child, while a parent whose formal schooling ended with the fourth grade is a negative influence. In the directional model, a "high father, low mother" combination is equated with a "low father, high mother" combination. In fact, however, quite different influences on a child's educational aspirations and achievements may flow from these two combinations. The results of the various combinations of parental educational level may vary among ethnic groups, occupations, and other variables as well.[14]

---

[13] This statement is subject, of course, to the usual qualification of "other things being equal," and also to the fineness of the definition of "similarity." Where definitions of similarity are too crude, interaction effects can produce differences between the members of a pair who are recorded as "alike" by the measures used.

We must keep clearly in mind, therefore, the possible risks in developing matching processes that permit the use of offsetting forces. But we are talking in terms of possible interaction effects. If they are, in fact, fairly unimportant, the stringent ideal model may actually obscure influences of the experimental variable. In this kind of problematic situation, it seems wise to make calculations using both models.

The research task in using the directional model is, of course, to separate the effects of mismatching from the effects of the experimental variable. We shall follow this procedure: Each experimental-control pair has been compared on the five critical variables. A zero is recorded when they are alike on a variable; +1 indicates that the experimental subject was higher (that is, he has a predisposition, or an advantage, in moving in the direction promoted by the stimulus); -1 indicates that the control subject was higher. The range of possible comparison scores, then, is from +5 to -5.

---

[14] Where statistical assumptions are met, inspection of correlation and regression plots may furnish us with checks on our directional ratings. Covariance or multiple regression analysis may reveal the single and joint effects of the complex of variables.

The pairs are also compared on an index of academic success. For simplicity in this explanatory section we shall assume that this index can also range from positive to negative, with the same meaning attached: +5 means a substantial contrast in favor of the experimental subject, and -5 a contrast in favor of the control subject. If we plot the correlations between the two sets of information for each pair, the location of the regression line can tell us whether the experimental stimulus produced the observed result, and, if it did, the extent to which it was supported by the predispositional influences. The lines in Figure 2 below indicate some of the possibilities. For purposes of calculation, the index numbers (in parenthesis) have been transposed into a positive series.

Regression line A would indicate that the experimental stimulus had no influence. Whatever advantage a person has over his partner to start with is precisely the advantage he ends with at the conclusion of the experimental period. Point x, for example, would represent a pair in which the control subject had stronger predispositions or supports for academic success at the beginning of the period, and at the end was found to have maintained the same advantage. Line B would indicate a situation in which the predispositions were more powerful than the experimental stimulus, but in which the latter produced some effect. Point y, for example, indicates that an experimental subject has reduced

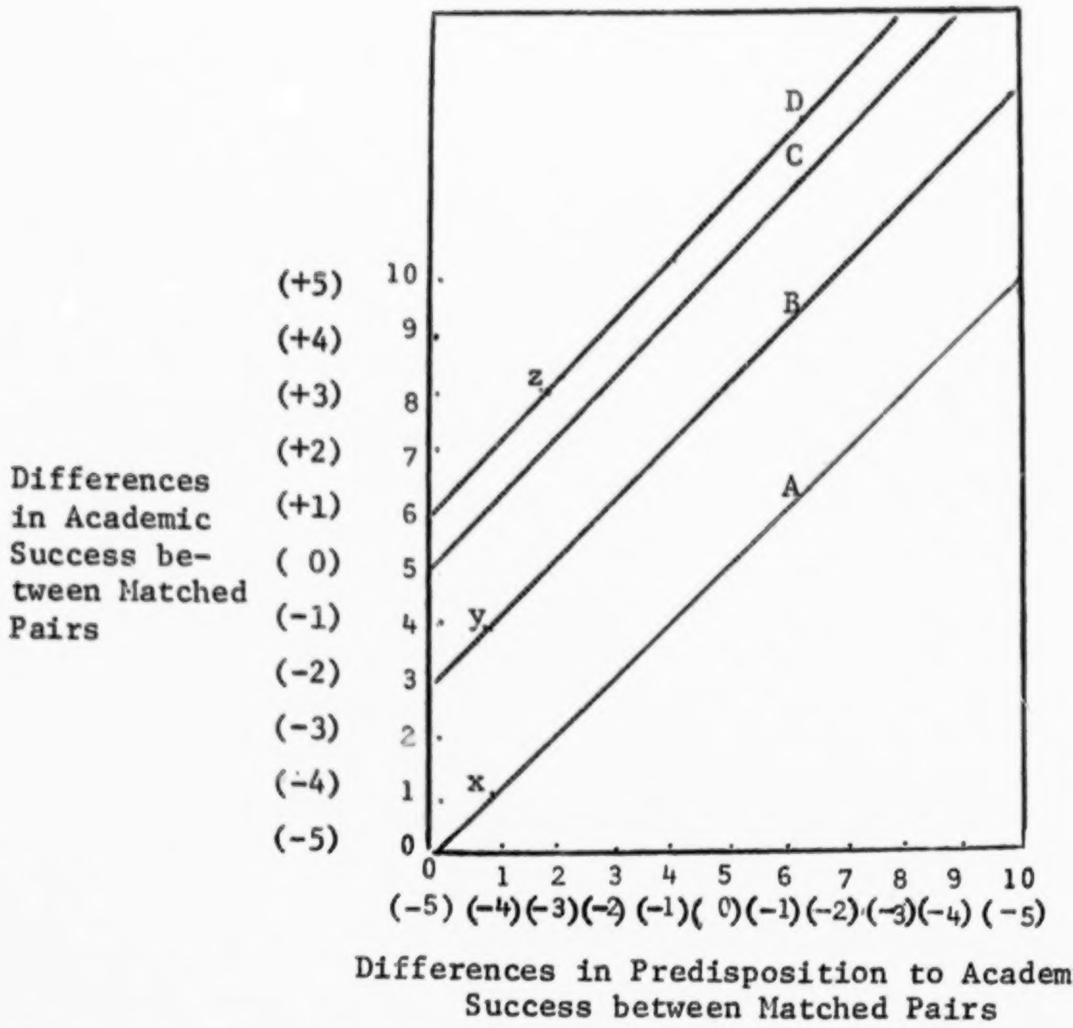

Figure 2.--Possible Correlations Between Differences in Predisposition and Differences in Academic Success Between Matched Pairs

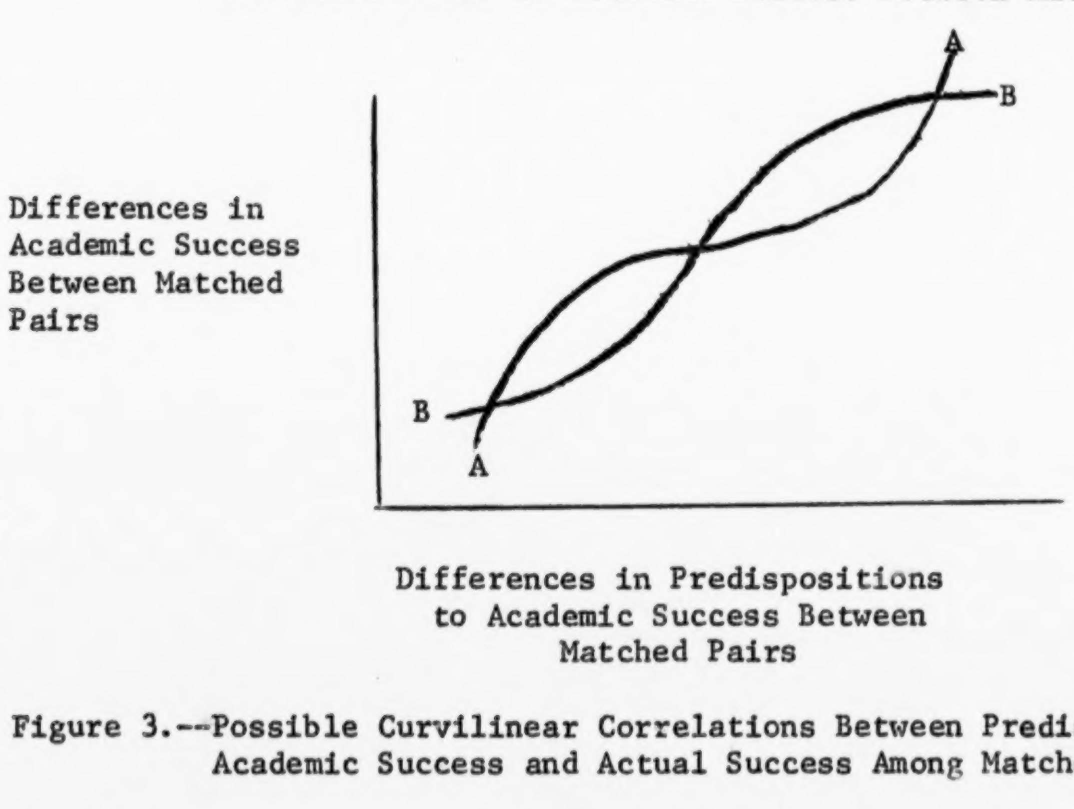

Figure 3.--Possible Curvilinear Correlations Between Predispositions to Academic Success and Actual Success Among Matched Pairs

an initial advantage of his control partner, but has not eliminated it. A regression line that took the pattern of C would demonstrate that initial advantages and the experimental stimulus were of equal strength. Line D would indicate that the stimulus more than compensated for an initial disadvantage. At point z, an initial disadvantage of the experimental subject (-3) would have been turned into a substantial advantage (+3).

Needless to say, there are many other possible relationships. If the experiment actually produces the opposite effect from that intended, the lines will slant from upper left to lower right. Or, more probably, curvilinear relationships may emerge, with an initial advantage having a progressively more powerful effect at the extremes (line A in Figure 3 below), or, oppositely, fading out at the extremes (line B).

Interpreting Possible Differences Produced By the Two Models. With two modes of comparison, based on the ideal and the directional matching indexes, we are confronted with the possibility that they may point to different conclusions. If they do, interpretations will be problematic, but this is preferable to a "certainty" based on insufficient analysis. There are three possibilities for each of the matching models: the experimental group may prove to be academically superior, no difference may be revealed, or the control group may be superior. When the two modes of matching models are used together, nine possibilities occur, as in Figure 4 below.

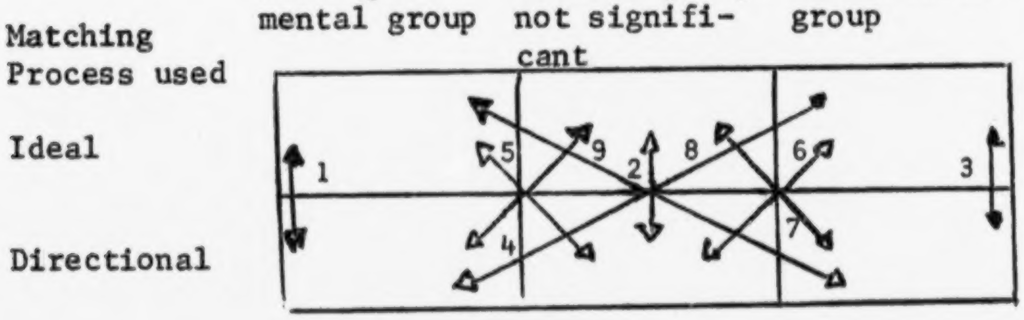
Advantage to the Experimental group
Differences between groups not significant
Advantage to the Control group
Matching Process used
Ideal
Directional
1
2
3
4
5
6
7
8
9

Figure 4

Each of these relationships presents substantial problems of interpretation, but we shall disregard here lines 6 through 9, and concentrate on the empirically more likely or theoretically more interesting possibilities. Line 1 would indicate that whether we compared the academic success of two persons who had been matched on many criteria or on five criteria only, with direction taken into account, the result was the same. The efficacy of the experimental program is supported.[15] Line 2 would indicate that the stimulus had made no difference, while Line 3 would point strongly to boomerang effects. Lines 4 and 5 are more problematic. An "efficacy" interpretation is supported by the use of "directional" matching data in line 4. This might be a result, however, of the fact that negative interactional effects were disregarded, for the ideal matching process reveals no significant differences. It may be possible to choose between these interpretations by a detailed comparison of pairs represented by lines 1 and 2 with those represented by line 4. Line 5 indicates a situation in which the ideal matching process supports an efficacy interpretation, while the directional matching process indicates no significant difference as a result of the experimental stimulus. The latter may be a result of the fact that positive or supportive interactional effects have not been taken into account, i.e., the match may be inadequate. Again, comparison of pairs represented by lines 1 and 2 with those represented by 5 may permit a choice between competing hypotheses.

---

[15] It may be well to note again the "arbitrary" (but not theoretically careless) quality of even the "ideal" matching scale. By employing scales of somewhat different form, with variation in weighting, we can discover whether the results obtained are influenced significantly by the scaling process.

The net effect of the operations we shall follow will not be to eliminate the possible distorting influences of inadequate matching. They will, however, allow us to study the effects of various levels of mismatching systematically, thus strengthening our interpretations. The use of random assignment to experimental and control groups supports the process. In addition, we shall use alternative procedures. By testing one procedure against another, we will be better able to determine whether or not our findings yield anything of significance, statistically and theoretically.

In sociological experiments, somewhat light-hearted references to to the fact that "subjects were matched on six critical variables" may be scarcely more than an expression of hope. On the other hand, we do not need to feel helpless in the face of the complexity of the research problem. While working for "identical" control and experimental groups, we can improve our analysis by judicious use of logical and mathematical procedures.

CLOSENESS OF MATCH BETWEEN EXPERIMENTAL-CONTROL PAIRS

We turn now to the data from our experiment. Table 1 indicates the distribution of matching scores for each of the three years. In some cases, information was lacking for one or both members of a pair, making it necessary either to eliminate that pair from the analysis (if less than 75 per cent of the information was available) or to estimate a score (when only a few items were missing). There are three possible ways to estimate a probable score for a given variable:

1) A person for whom a measure is lacking can be assigned the midpoint of the range for that variable. Thus stanine 5 would be used,

in variables 14 and 15, to estimate probable scores for those individuals for whom intelligence and achievement test scores are lacking. Or, where the score for a variable is directly comparative, as in item 2, the midpoint is assigned to the pair, without reference to individual measures. If we had no information on a pair or on a variable, this "midpoint" method of estimation would be reasonable. We do have information, however, so that other modes of estimation are wiser. We know that the midpoint is not the best estimate of score for a variable with reference to all those pairs on whom information is available. For example, most pairs have a score of 0 for variable 1. We would distort the matching picture if we assigned a score of 2 (the midpoint on variable 1) when, in fact, the mode is 0 and the mean score less than .5.

2) Problems associated with the first method of estimating missing scores lead us to the second method: Assign to any pair for whom information is missing on a variable the mean score on that variable among all the pairs for whom information is complete. Thus if the mean contrast in religious interest of parents is .5, as calculated in variable 18, that score would be assigned to a pair for whom information is missing, rather than 1.0, which is the midpoint score. This procedure is based on the assumption that in a relatively homogeneous group, known scores are a better indicator than the midpoint of an arbitrary range.

3) The third method of estimation extrapolates from what is known about a given pair to fill in missing information. If, for example, we have data on 17 variables for a pair, with an index of congruence of 16 points out of a total possible 64 points, we can assign a score of 1 for the missing variable, which has a possible score of 4. In using the same

ratio (1/4 : 16/64), we have assumed that existing information about a pair is the best source of estimation of missing data. This procedure treats the unmeasured variable neutrally, not letting it affect the observed relationship based on the measured variables, but making it unnecessary to remove a pair from the analysis because of the missing items.

For the most part we shall use the third method of data estimation. Clearly there are risks involved in any one of the procedures suggested; but in our judgment the third method is least likely to over-or-under-estimate the quality of the match. To check against the possibility that this procedure influences the results, however, we shall compare those pairs who were most fully matched with those least fully matched. If our hypothesis is significantly more strongly supported in one group than in the other, we shall conclude that the process of estimation has influenced the results.

Table 1

Distribution of Matching Scores
Ideal Index
(Point differences stated as percentages of points possible)

Per Cent Difference	1964 (n=56)	1965 (n=70)	1966 (n=69)	Total (n=195)
0.0-9.99	7	9	10	26
10.0-19.99	24	33	26	83
20.0-29.99	15	22	26	63
30.0-above	10	6	7	23
mean %	20.07	18.18	18.60	18.87

The data in Table 1 indicate that our matches, using an 18-variable scale, were imprecise. When set against what we have called an "ideal" match, at any rate, our pairs show important differences. The average mis-match on the 68 point scale was approximately 13 (18.87 per cent). (We used percentages to simplify comparison among pairs with varying amounts of information.) Without some standard of comparison, we cannot say that this represents good or poor matching. And since this is a new procedure, there is no such standard. It is our judgment, however, that in this kind of field research, with the inevitable constraints among five school systems, it would be difficult to get average matching scores that fell much below the 18 percent level.

The directional matching process yields more similar pairs:

Table 2

Distribution of Matching Scores
Directional Index

		1964 (n=56)	1965 (n=70)	1966 (n=69)	Total (n=195)
Maximum advantage to experimental child	+5	1	2	1	4
	+4	0	1	2	3
	+3	9	3	6	18
	+2	11	10	14	35
	+1	12	11	10	33
Equality	0	6	17	9	32
Maximum advantage to control child	-1	4	14	7	25
	-2	7	5	10	22
	-3	5	6	6	17
	-4	1	0	4	5
	-5	0	1	0	1
	mean	0.52	0.10	0.12	0.23

Although Table 1 shows that the matches, using 18 variables, were imprecise, Table 2 shows, as we expected, that there was no large advantage to either experimental or control groups. Using those criteria (family support and pre-program tests and grades) that we hypothesized would give a child an advantage,[16] we find that the experimental group is slightly better off, particularly in the 1964 cohort, but that the advantage is small. The correlation between the ideal matching score and the directional matching score is very low — .04, indicating that our randomizing procedures produced groups of nearly equal potential for academic success.

We are primarily interested in comparing individual pairs, however, and not group means, hence it is important to notice the average distance from parity, disregarding signs. The average advantage of one or the other of the members of a pair, (slightly more likely to be the experimental member of the pair as shown in Table 2) was as follows: 1964 = 1.84; 1965 = 1.44; 1966 = 1.88; Total = 1.71. These differences are based on a possible advantage score of 5.0.

These measures of initial advantage or equality need to be seen in connection with measures of the post-program levels of achievement. We shall measure the latter by means of an index composed of six items (maximum weights are given in parentheses): Number of post-program years of school completed (6); shift to a higher "track" in school (1); move to a private or special academic school (1); achievement test scores (we used stanine differences, where available, with a maximum possible contrast of 8, but with an empirical range of 4; where stanines were not available, we used test-years, with each one-half year contrast equalling one point); junior high school grades (3); senior high school grades (3).[17] Without reference to initial advantage of the or the other member of a pair, the distribution of dependent variable index scores was as follows:

---

[16] 16. See Postscript at end of this chapter for procedures used in calculating the Directional Matching Index.

Table 3

Comparison of Post-Program Achievement Levels of Matched Pairs

Dependent Variable Index	n	
+10	4	
9	4	
8	7	
7	8	
6	12	Total positive (experimental advantage) = 118
5	15	
4	14	
3	14	
2	19	
+1	21	
0	21	
-1	10	
2	13	
3	15	
4	9	
5	4	
6	2	Total Negative (control advantage) = 56
7	1	
8	2	
9	0	
-10	0	

---

[17] See the postscript at the end of this chapter for procedures used in calculating the Dependent Variable Index.

We shall not be able to test the influence of the program unless we consider the initial advantage to one or another child. This can be done in a variety of ways. If we divide the columns in Table 2 between +1 and 0 we find that 93 of the experimental children had a pre-program advantage, and 102 of the control children had such an advantage or equality. Table 3 applies a simple chi-square test to the null hypothesis: the distribution of advantage scores did not shift significantly as a result of the program; those members of each pair who had higher advantage scores before the program retained their advantage after the program. Original advantage scores are used for the expected frequencies.

Table 4

Shift of Advantage Score as a Result of the Program

	Experimental group		Control group	
	observed (t_2)	expected (t_1)	observed (t_2)	expected (t_1)
Later advantage score	118	93	77	102
Later disadvantage score	77	102	118	93

$X^2 = 25.70$ p. = $<.001$

[d.f. = 1]

t_1 = before the summer program
t_2 = after the program

By the chi-square test, the experimental subjects significantly improved their position in relationship to their partners, as measured by the index of academic achievement. The chi-square test, using dichotomous variables, may hide variations within each quadrant, however, even though the total relationship significantly favors the experimental group. As a further test, therefore, we can break down the independent variable (original advantage from family background, pre-program test scores and school performance) and the dependent variable (post-program levels of achievement) into more refined categories. Table 4 shows "steps" of gain or loss vis-à-vis one's partner, the maximum being +4 or -4, with positive numbers indicating a gain for the experimental child and negative numbers indicating a gain for the control child. (Because of the small numbers at the extremes in the two indexes, we have consolidated the scales into 2 to -2 ranges.)

Table 5

Pre-Program to Post-Program Gain or Loss in

Academic Performance between Matched Pairs

		Advantage score on index of academic achievement					
		-2	-1	0	+1	+2	Total
Advantage score on directional index	+2	0	1	1	1	5	8
	+1	0	10	8	26	9	53
	0	2	16	28	28	16	90
	-1	2	10	14	7	5	39
	-2	1	4	1	0	0	6
Total		5	41	52	62	35	195

The diagonal scores (in boxes) indicate the pairs who did not change in relationship to one another. In the lower left, for example, one control child started out with a two-step advantage over his partner and retains, by our measures, a two-step advantage. Each cell can be assigned a "score" by reference to the two indexes. Thus in the "-1 and +2" cell on the lower right there are five pairs where the experimental child started out with a one-step disadvantage (-1), but now holds a two-step advantage (+2), a gain of three steps.

Whether we cite the number of individuals who improved their position or the number of steps by which positions were improved, the results match those of the chi-square test; they also reveal small additional gains by some experimental children that were obscured in the 2x2 table. The data in Table 5 indicate that there was no change in the relative scores of 70 pairs; in 84 pairs, the experimental child gained (a total of 118 scale steps); and 41 of the control children gained (a total of 56 scale steps). To put it in other terms, among the experimental children, 43% gained, 36% held their own, and 21% lost, relative to their partners.

Various measures of correlation can help us to interpret the effects of the program. We have estimated missing information on a variable by assigning to that variable the mean percentage difference between the members of a pair that was found on measured variables. Did this estimation produce an apparent effect from the program where there was no such effect? Apparently not. The higher the matching percentage (that is, the more information available), the lower the score on the dependent variable index (high score = advantage to the experimental child), but the correlation is low and not significant (-.04). Advantage to the experimental child is not a function of poorer information.

There is also a correlation between matching score on the ideal index (indicating closeness of match, not amount of information), and the dependent variable index: r = .11. That is, the more different the members of a pair were to begin with (without reference to advantage to one child or the other, since we are referring here to the ideal, not the directional index), the more likely it was that the experimental child would show an advantage after the program. Referring back to Figure 1, we can note that the regression line, although closer to line B than line A, is not horizontal. However, the correlation is not significantly higher than zero.

It should be noted that the correlation between the directional match and the dependent variable index is .30. The variables included in the directional index are part of the ideal index, and account for the relationship between the quality of the original match and the measured levels of academic achievement. As we have noted above, if a child started out with an advantage, he often ended up with an advantage (as shown by the correlation of .30). In nearly two-thirds (64%) of the cases, however, this did not happen. Both experimental and control children gained, but the former twice as frequently as the latter.

POSTSCRIPT

I. Procedures for determining the Directional Matching Index. Combine scores on the following items:

1. Composite score on father's status

A. Educational comparison (see variable 8 Index of Congruence)	B. Occupational comparison (see variable 7)	C. Status in home (see variable 6)
+2 = Ex 3 or 4 higher	+2 = Ex 3 or 4 higher	+2 = Ex 3 or more higher
+1 = Ex 1 or 2 higher	+1 = Ex 1 or 2 higher	+1 = Ex 1 or 2 higher
0 = educ. equivalence	0 = Occ. equivalence	0 = equivalence
-1 = C 1 or 2 higher	-1 = C 1 or 2 higher	-1 = C 1 or 2 higher
-2 = C 3 or 4 higher	-2 = C 3 or 4 higher	-2 = C 3 or more higher

2. Composite score on mother's status

A. Educational comparison	B. Occupational comparison I. Status (see variable 10)	B. Occupational comparison II. Skill (see variable 11)	C. Status in home
(as above)	+1 = Ex 2 or 3 C 0 or 1	+1 = Ex 2 or 3 C 0 or 1	(as above)
	0 = Both 2 or 3 or 0 or 1	0 = Both 2 or 3 or 0 or 1	
	-1 = C 2 or 3 Ex 0 or 1	-1 = C 2 or 3 Ex 0 or 1	

These operations yield possible scores ranging from +6 to -6 on variables 1 and 2. Reduce each to a four point contrast in the following way:

+6 +5	= +2
+4 +3 +2	= +1
+1 0 -1	= 0
-2 -3	= -1
-4 -5 -6	= -2

3. Intelligence tests (sixth or seventh grade)

+2 = Ex 2 or more stanines ahead
+1 = Ex 1 stanine ahead
0 = same stanine
-1 = C 1 stanine ahead
-2 = C 2 or more stanines ahead

4. Achievement tests (sixth or seventh grade)

(as above, in 3)

5. Grades in seventh grade "solid" subjects (English, Science, Mathematics, Social Studies), average on 0-4 point scale

+2 = Ex 1.00 or more ahead
+1 = Ex .50-.99 ahead
0 = neither more than .49 ahead
-1 = C .50-.99 ahead
-2 = C 1.00 or more ahead

These operations yield possible Directional Matching Index scores ranging from +10 to -10. Since only two of the 195 pairs, however, were outside a +5 to -5 range, we have consolidated into a +5 to -5 scale.

II. Procedures for determining the Dependent Variable Index.

Combine scores on the following items, using positive sign when advantage is with the experimental child, negative sign when advantage is with the control child:

1. School achievement level. (Score is determined by point difference between experimental and control subjects)

    1. Less than eleventh grade completed
    2. Eleventh grade or above, but not graduated from high school
    3. High school graduate
    4. Post high school training (technical or vocational school)
    5. Junior or community college
    6. Four year college or university
    7. Select four year college or university

2. Special track in high school.

    0. Less than half of school years after Special Opportunity Program in special track
    1. One-half or more of post-S.O.P. years were in special track

3. Special school.

    0. Not in special school after Special Opportunity Program
    1. In special school (private or special academic school)

4. Achievement tests.

    Where stanine scores are available, record difference in stanines.
    Where stanines are not available, use following measures:
    0. Less than half year difference in achievement test scores
    1. One-half to .99 years difference
    2. One to 1.49 years difference
    3. 1.5 to 1.99 years difference
    4. 2.0 to 2.49 years difference
        (To match the possible range of stanines, that is +8 to -8 this scale continues to maximum of 4.5 years difference; but in fact, only one pair showed more than a 2.5 years contrast.)

5. Junior High School grades (after the S.O.P. program), solid subjects.

    0. Less than .50 difference in average grade
    1. .50 to .99 difference
    2. 1.00 to 1.49 difference
    3. 1.50 or above average difference

6. Senior High School grades, solid subjects (Grades 10 through 12)

    (Calculate as in variable 5)

Chapter Five: SIGNIFICANT OTHERS AND EDUCATIONAL ATTAINMENT

The ability of a child to attain well in educational and occupational terms depends to a significant extent upon both his talent/capacity and on the socio-economic status of the parents. The review of the literature in Chapter 1 suggests the importance of these two distinct influences on attainment.[1] For a child to actively pursue in high school an academic and personal career that would merit admission to college or university, however, additional influences may be involved. These bear on the psychological outlook of the child (his aspirations, plans, and choices) as well as on social and psychological interaction within a network of persons and groups of persons. For the child, the development of specific expectations and plans, of knowledge of resources to continue upward attainments, of a favorable self-image and associated motives and drives of ambition or achievement to obtain a sense of mastery in the face of extreme difficulties and obstacles, depend upon a felt sense of support and guidance involving a complex network of persons and groups.[2]

---

[1] See Theodore D. Kemper, "Reference Groups, Socialization, and Achievement," AMERICAN SOCIOLOGICAL REVIEW, February, 1968, pp. 31-45, for a most recent conceptual treatment. See also Theodore L. Reed, "Reference Groups and Educational Achievement," unpub. M. A., Oberlin College, 1966; Margaret Rhile, "Social Class Origins and Interpersonal Supports for Educational Mobility," unpub. ms., University of Wisconsin, 1968; and Susan Margolis, "Significant Others and Disadvantaged Youths," unpub. ms., Oberlin College, 1964, on our efforts to organize the theory and methods for study in this area. Cf., Robert E. Herriott, "Some Social Determinants of Educational Aspiration," HARVARD EDUCATION REVIEW, Spring, 1963, pp. 157-177; Ralph Turner, THE SOCIAL CONTEXT OF AMBITION, Chandler, 1964, pp. 1-18, 218-255; Ernest Q. Campbell, and C. Norman Alexander, "Structural Effects and Interpersonal Relationships," AMERICAN JOURNAL OF SOCIOLOGY, November, 1965, pp. 284-289; Eugene Litwak and Ivan Szelenyi, "Primary Group Structures and Their Functions: Kin, Neighbors, and Friends," AMERICAN SOCIOLOGICAL REVIEW, August, 1969, pp. 465-481. See also Eugene Litwak and J. Figuera, "Technological Innovation and Theoretical Functions of Primary Group and Bureaucratic Structures," AMERICAN JOURNAL OF SOCIOLOGY, 1968, pp. 468-471.

On summary findings on the relative power of parents and peers in influencing adolescent choices in educational and personal plans, see Denies B. Kandel and Gerald B. Lesser, "Parental and Peer Influences on Educational Plans of Adolescents," AMERICAN SOCIOLOGICAL REVIEW, 34, April, 1969, pp. 213-223. Cf., George Y. M. Won, Douglas S. Yamamura, and Kiyoshi Ikeda, "The Relation of Communication with Parents and Peers to Deviant Behavior of Youths," JOURNAL OF MARRIAGE AND FAMILY LIVING, 31 (February, 1969), 43-47.

As these social-psychological inputs have been added to the initial analyses, which included measures of capacity and measures of parental status, important variation has been accounted for in the data on educational and occupational attainment. We have sought to measure the influence of "significant others" on the direction and quality of educational attainment of the children in the study.

> Significant others are the specific persons from whom the individual obtains his level of aspiration, either because they serve as models or because they communicate to him

---

[2] See the review by Rhile on this point, op. cit., passim. See Richard L. Simpson, "Mobility Orientation and Stratification," AMERICAN SOCIOLOGICAL REVIEW 22, 1957, pp. 204-212; Edward McDill and James Coleman, "Family and Peer Influences in College Plans of High School Students," SOCIOLOGY OF EDUCATION, 38, 1965, pp. 112-126; and William H. Sewell and Vimal P. Shah, "Social Class, Parental Encouragement and Educational Aspiration," AMERICAN JOURNAL OF SOCIOLOGY 73, March, 1968, 550-572. Cf., Herriott, op. cit.

Reed finds some tendency among "less advantaged" junior high school youths to have established "outside" contacts with adults and/or peers who are "more advantaged," and thus probably able to orient more towards high educational achievement. Rhile also finds "compensatory" sponsorship of a more direct sort among liberal arts college students. Note also Edward Laumann's PRESTIGE AND ASSOCIATION IN AN URBAN COMMUNITY (New York: Bobbs-Merrill, 1966), on the "blurring" effect among upper working class and lower middle class groupings. The fact that majorities may promote values which turn away from dominant achievement and mobility values is described by Herbert J. Gans in his URBAN VILLAGERS (Glencoe: Free Press, 1962).

> their expectations for his behavior. The term "significant other" is more appropriate than "reference group" because it eliminates the implication that collectivities such as one's friends, or work groups, or parents are necessarily the influential agents for all individuals.[3]

Based on a review of both the theoretical and empirical literature, we developed measures to index the influence of individuals who represented valued educational and/or occupational attainment in one way or another.

These networks form different configurations for different families and so are a logically possible source for explaining the variation between families. They include those individuals who interact regularly with the family as well as those who are perceived by the family to be important to them, who may or may not engage in actual interaction with them. Thus an aunt who has been to college and teaches school may be an important significant other to a particular family even though they may see her only three or four times a year. President Kennedy, who showed that the son of an immigrant can make it, might be a significant other in the life space of another family whom he had never known. Generally speaking, however, significant others are those persons the family sees as most important and/or most congenial within the kin and peer/friendship groups.

Significant others often symbolize one or more facets of mobility and achievement. They provide living examples of the consequences of achievements (or non-achievements); standards for comparing one's position in life to his potential position, examples of how a particular

---

3. William H. Sewell, Archibald O. Haller and Alejandro Portes, "The Educational and Early Occupational Attainment Process: Wisconsin Farm-Reared Men, (1957-1964)," AMERICAN SOCIOLOGICAL REVIEW 34, February, 1969, 82-92, p. 83.

standard is achieved, and, perhaps, a general incentive to action. Those significant others in regular interaction with the family can be a conscious and powerful influence on the family's social mobility and achievements. Even where a significant other does not consciously try to influence a family's movements, his support and approval can be crucial for family decision making.

Significant others are often counted on for the following kinds of direct supports:

1. Material Supports: Persons and families often interact with those close to them in financial and non-financial exchanges to help the recipient in pursuing a given valued activity. Where large expenditures are involved, a "mutual aid fund" is often available from within the informal kin and family network, occasionally coming almost entirely from one particular member.

2. Moral and Emotional Support: The significant other network often provides moral or social validation for a person or family's past, present, and contemplated actions. The network, or important members within it, may also **disapprove** of such actions; and such disapproval may carry great weight. The group of significant others may be crucial in determining whether a particular achievement effort is greeted with approbation or ridicule.

3. Instruction: Significant others often provide overt instruction for individuals and families in how to think as well as how to behave. Probably just as important is the covert instruction provided by those members of the significant other network whom an individual or family

adopts as role models. Significant others may by word or example teach families how to utilize available means for achievement.[4]

4. Advocacy: A fourth major function of the significant other network is their direct or indirect participation in family life by serving as advocates--those who will intercede for the individual and the family in their achievement efforts. One example is the relative whose recommendation secures an individual a particular job. Another is a family friend who knows enough about the school system to intercede and see to it that the family's bright child is placed in the college-oriented curriculum.

The influence of significant others extends beyond the direct functions they fill. Even where they do not provide tangible aid of the t types described above, significant others may be important because:

(1) Influencing behavior is often informal, sincere, and intrinsically rewarding. Think of praise coming from a valued relative.

(2) Significant others are often present when a family makes small but important advances and thus can provide immediate reinforcement.

(3) One's evaluation of reality is influenced by the judgment of his peers. Significant others may be important in shaping one's perceptions of the objective possibility of advancing along particular lines; they thus may affect the probability that one will in fact try to advance along those lines.

---

[4] See Frank W. Young and Ruth C. Young, "Invidivual Commitment to Industrialization in Rural Mexico," AMERICAN JOURNAL OF SOCIOLOGY 71, January, 1966, pp. 373-383; see also Elihu Katz, "The Two-Step Flow of Communication: An Up-to-Date Report of an Hypothesis," PUBLIC OPINION QUARTERLY 21, Spring, 1957, pp. 61-78.

(4) Achievements by one in some ways regarded as an equal may convince an individual that he has the potential for similar achievements.

(5) Symbolic association with one who has made it may lead to a real rise in a family's status and an actual improvement in a family's life chances even in the absence of any other aid from the significant other.

(6) Some of the social-psychological literature suggests that one individual expressing a "minority opinion" can significantly strengthen the convictions of one tending toward the same "minority opinion." If this holds outside of experimental laboratory settings, then one person who thinks going to college is important may be able to offer effective psychological supports to another who leans in that direction even though the majority of the latter's adult and peer relations think high education is either a silly or non-attainable goal. It is expected that such influence would increase with the prestige status of the individual holding the supporting opinion. Although it seems reasonable that the more individuals a person has to support his personal opinion, the stronger their probable influence, yet two persons at the maximum seem sufficient in laboratory studies. Minority opinions can thus be held in the face of large majorities in the opposite direction.

(7) An individual's continual interaction with a set of significant others may lead to a "generalized other" or a sense of conscience which guides responses to small cues when supporting others are not present. The person generalizes from past experience with significant others, their ideas, and their behavior so that cues in the individual's environment, which would have been unfamiliar or absent in previous interaction,

serve to remind him of what is necessary for continuing achievement. Thus self-image and related motives and drives may be developed to enable an individual to sustain a given line of achievement with minimal external rewards and stimulation over an extended period.

From this perspective, relative deprivation and "poverty" depend upon more than relative income and general socio-economic status. The cynical statement, "It's not what you know, but who you know," becomes true, not because what one knows is unimportant, but because who he knows may be a necessary precondition for the acquisition of valuable knowledge. Those families whose significant other networks include functionaries with expertise on the methods of mobility (e.g., teachers of social workers) may have a special advantage even if they are unable to get any special favors because of the relationship. Where one has access to significant others who can give special favors or financial help, the advantage may be even greater.[5]

Members of the low income population differ markedly in the quality and extent of their significant other networks. Some feel close to few individuals and others feel close to many. Some can turn to many in emergency situations; others to none. And some have friends or relatives who are successes by almost any definition, while others are close only to those whom society would style "failures."

The family with a large number of unsuccessful significant others may be worse off, from the standpoint of its mobility prospects, than a family with almost none held in esteem. Many of the same mechanisms which allow significant others to advance the prospects of a family mobility may cause the significant other network to impede such advancement. Significant others have the opportunity not only to aid, but demand aid. If they can maintain a legitimated right to insist on extra supports from the most successful member of their network, they may prevent that member from ever leaving low income status. They may punish mobility attempts as well as reward them. And they may teach and value skills, such as certain criminal skills, which are functional for certain short term goals but fatal to long term mobility prospects.

---

[5] These interactive effects (see Chapter 1) are noted by Kemper, op. cit., pp. 40-42, under the concept of a coincidence of functions.

Occasionally, when significant others who could aid in mobility are lacking in the family's acquaintance network, a third party, who initially has no friendship or kin ties with the family, may be led into interaction with the family by his role demands and may come to fulfill the functions of a significant other without every establishing a viable peer relationship. One example of this is the teacher who takes a particularly promising child under her wing and becomes the role model for the child in his future performance. Such individuals may also play a crucial role when the significant other network is generally supportive but lacks the knowledges or resources to aid a particular family in its mobility efforts.

The Participation of Administrative-Legal Functionaries as Significant Others. Administrative and legal functionaries can be important as significant others:

> Clearly, the variable we have called significant other influence is an important factor. The present evidence appears to show that once formed, its effects are far-reaching. Also, besides being a powerful explanatory factor, it should be amenable to manipulation (in line with public policy). It thus suggests itself as a point

> at which external agents might intervene to change educational and occupational attainment levels. This means that at least part of the system is theoretically amenable to experimental testing. It also may mean that practical change agents might be able to change levels of attainment, either by inserting themselves or others as new significant others or by changing the expectations existing significant others have for the individual. There may well be a substantial pay-off from more refined work with this variable.[6]

Here, an important policy question is revealed. If an influential social network provides the kind of unwitting support and guidance which orients family members in a given direction, can the staff involved in public and quasi-public interventions become part of this network through participation of a more intimate sort? If it cannot become part of the family's intimate circle, can it nonetheless make a difference in the likelihood that these families can encourage higher attainments among its children?

Personnel responsible to a hierarchical, formal organization are unlikely to become part of this intimate network. Schools and school personnel, nevertheless, significantly affect the educational opportunities of a child. Without becoming part of the network of significant others, these individuals affect the child's education by the continual testing for ability and the subsequent tracking of a child in given curriculum opportunities which lead to college admissions or deter such chances, by the sanctions employed to obtain compliance to school expectations, and by other indirect influences. In this sense, the effects of school programming or external educational interventions, such as the Special Opportunities Program, can be independent of the influence of significant others. If such administrative and legal interventions can be supported by significant others, however, or if the intervening staff can become part of this more intimate network, then the achievement or acquisition of skills and motives necessary for achievement should be more rapid. In either case, the chances for achievement should increase.

---

[6] Sewell, Haller and Portes, _op_. _cit_., p. 91.

Administrative and legal arrangements could be structured to provide educational services that participants would view as helpful in their efforts to attain educational goals. Indeed, there is a recursive effect. Once a family has decided that college education is an achievable goal, suggestions from external authority which earlier would have been viewed as impersonal may be accepted as "good advice." Pressures to maintain good grades, to control troublesome school behavior, to pay attention to routine but critical steps in test-taking and form-filling activities on schedule, may be viewed more favorably by the child and the family. In the case of the Special Opportunities Program, the moral-emotional support of its staff, the information and instruction it provides, and the intercession and advocacy the staff could make and does make to affected schools and agencies can become part of the family's own efforts at self-improvement. Thus, even without entering completely into the intimate network of significant others' public agents may perform equivalent functions when the goals of the families and the goals of the administrative body in fact converge. Where a family lacks initially coincident goals, however, and does not adopt the goals of the intervening body, the administrative-legal influences, even though powerful, will usually be viewed as imposed and external to family efforts.[7]

---

[7] See Jean Piaget, THE MORAL JUDGMENT OF THE CHILD (New York: Collier Books, 1963), Marjorie Gabain, trans., pp. 395-406 on suggestive qualities of "cooperation" in making the "moral rules" part of one's own efforts over rules made by hierarchical, heteronomous dissemination.

The intervention of relatively impersonal functionaries in the achievement process (as in education) can become "significant" to the extent that family members define the participation by the functionaries as part of a _self-chosen_ achievement effort. Coleman makes this point about the survival of individuals in an urban environment, when he talks about discretion and self-action becoming the basis for further efforts at achievements.[8] A family aided by an intervening body, such as S.O.P. can think of achievement as being of their own choosing in the same way that they can think of aid from a relative or close friend. _Objectively_, little advance is achieved _alone_ because those responsible for entry into, exit from, and evaluation of performance within a given status determine who will receive the necessary and extraordinary benefits and under what conditions such assistance will continue. The _theoretical_ problem is to determine how a normally "external influence" becomes more a matter of _internal_ choice. If we knew this, public services, such as in education, could be structured to become part of the self-chosen effort of the family and the larger community.

In principle, the Special Opportunities Program, through staffing and services, has attempted to become part of the family's self-chosen effort to aid one of its children, who was deemed "bright with potential for college." By direct probes and indirect questioning, we have obtained from our respondents both the indication of specific individuals and also indexes of their "power" in setting standards and providing diverse kinds of supports for higher attainments. It remains an empirical question to determine the degree to which the staff has come to be salient in the significant other network and whether the total effect has been to improve the child's chances for higher education.

---

[8] James S. Coleman, "Implications of the Findings on Alienation," AMERICAN JOURNAL OF SOCIOLOGY, 70 (July, 1964), pp. 66-68, 77-78. The letter comments on Melvin Seeman's article, "Alienation add Social Learning in a Reformatory," AMERICAN JOURNAL OF SOCIOLOGY, 69 (November, 1963), pp. 270-284.

"the most important element for survival in an urban society is how to take responsibility for one's self. . . it seems explicit educational policy in this direction could aid as well (as learning from the urban conditions themselves) . . . this can be taught by making the environment one in which the responses are eontingent upon the individual's own behavior. . . ," Coleman, p. 78, _ibid_.

Factors in the Influence of Individuals. The following factors appear to be important in affecting the ability of particular individuals to provide standards of educational conduct and to furnish direct and indirect instruction into the means and resources necessary for educational attainment:

1. The Prestige of the Individual. In a social setting where there are fewer highly achieving others than in higher socio-economic status settings, the availability of some individuals with high prestige can be critical. This influence is indexed by a description of the relative status of the individuals mentioned by educational and occupational levels combined. Only those with high educational attainment, with variation in occupational attainment, are counted in this analysis.

2. Relationship in Social Structure. The relationship of the given individual of high attainment to the target child is described by specific social distance and by more general breaks in social distancce (kin network, friend network, and school and public officials network). In general, there is a greater advantage to having a person of high attainment in the kin network over the friend network over the public functionaries. Mention of S.O.P. staff (as public functionaries) is reported in this section.

3. Social-Psychological Function. Each person mentioned is described in terms of being close and/or admired by the adult and/or the target child. Where both share in selected social-psychological ties, the power of that person is assumed to be greater. A shared coalition effect would be greater than where only the mother or only the child identifies a person as being close or admirable.

4. Household Presence. If an individual ever lived or lives presently in the immediate household of the target child, then, that individual has additional power to influence through role modeling and accessibility for instructional purposes.

For preliminary analysis here, significant others are defined as those individuals who are cited by the mothers or female guardians as close to and/or looked up to by herself. These persons have either attended some college and/or are semi-professional or above in occupational status. The range of citations reported by the mothers vary from none to 28 persons, with a mean of 3.1 citations.

FINDINGS IN CONNECTION WITH SIGNIFICANT OTHERS

Three alternative hypotheses are suggested with regard to the role of significant others in influencing the educational plans and achievements of the participants in the study. In terms of program efforts, a mediational hypothesis is proposed. During the initial summer program and in subsequent follow-up activities, well-achieving persons from the same social-ethnic backgrounds as those of the students and their families were invited to speak about significant educational and occupational careers. The hope was that these presentations would lead the child and the family involved to seek out persons in their own social and reference group networks who might aid them in further educational/occupational planning and achievements.

Whether program efforts succeeded or not, given families may have begun with a "thinner" or "thicker" network of significant others before talent search and selection was completed. Those with a greater number of accessible and well-achieving persons visible to the family and to the child would be expected to continue to do much better than those whose significant other networks were "thinner." This hypothesis would stress an overriding pre-program influence, based on initial differences between the matched pairs.

A third hypothesis would combine the influence of initial differences and program effects, such that both influences would be significant either in enabling a student to do well or in limiting his achievements and plans. The findings reported below may provide a basis for evaluating these three alternative hypotheses.

Before and After Measures of Program Efforts. For each matched pair in the study, we first classified the experimental child as either beginning below the control child, similar to the control child, or above the control child in initial educational resources (exclusive of access to significant others). For each matched pair, we then classified the experimental child as below, similar to, or above the control child in the dependent variable or the index of educational achievement.

Table 5.1 reports parallel findings on the impact of the program upon experimental children classified initially as below, similar to, and above the control child in educational resources. The first set of findings concerns the relative rank of the experimental child on an index of educational achievement at the end of the study as compared with his matched control. Whether we examine the information for the full

Table 5.1

Relative Standing on the Index of Educational Achievement
By Initial Level of Educational Advantage and
Program Participation

Subsequent Standing on Index of Educational Achievement	Initial Level of Educational Advantage/Disadvantage of Experimental Child			Total
	Resources Below Control Child (-)	Resources Similar to Control Child (=)	Resources Above Control Child (+)	
Achievement above that of Control Child (+)	33% (37%)	50% (54%)	63% (61%)	51% (52%)
Achievement similar to that of Control Child (=)	33% (30%)	30% (18%)	20% (20%)	26% (22%)
Achievement below that of Control Child (-)	33% (33%)	20% (28%)	17% (20%)	23% (26%)
Total	99%* (100%)	100% (100%)	100% (101%)*	100% (100%)
N (Total matched pairs)	45	90	60	195
N^1(Matched pairs with information on Significant Others)	30	50	41	121

195 matched pairs or the 121 matched pairs for whom solid information about significant others could be obtained from the interviews, the pattern of findings is similar:

1. About 2/3 of the experimental children who initially began lower in educational advantage are higher in the subsequent level of educational achievement at the end of the study, (one-third in this category "lose ground.")
2. About 1/2 of the program participants who began at a similar level of educational resources end up ahead of the control child, (one out of 5 loses ground).
3. About 2/3 of the program participants who began ahead of the control child maintain that difference in educational achievement at the end of the study (one out of 5 loses ground).

This can be put in one figure, as noted in Chapter Four, by stating that 43 per cent of the experimental children moved ahead of their counterpart controls, while 21 per cent fell behind, with the rest remaining equal.

ACCESS TO SIGNIFICANT OTHERS

Each matched pair was studied to determine which member of the pair had more or less access to significant others. The mean difference between the experimental and control groups was .5, with 3.4 mentions and 2.9 mentions respectively. The pair by pair comparisons, summarized below, suggest systematic variation in access to significant others in relation to the relative level of educational resources and educational achievement among the pairs.

The terminal differences in relative educational achievement can be classified on three levels: that of the experimental child who is below the control child, equal to the control child, or above the control child in achievement. As we have noted, each of these three levels can, in turn, be subdivided on the basis of pre-program advantage (+) disadvantage (-), or equality (=). Our concern here is to specify the

Table 5.2
Mean Differences and Pair-by-Pair Differences in Access to Significant Others by Program Participation and by Resource/Achievement Level
(N = 242, 121 pairs)

Resource/Achievement Profile of Experimental Child	Mean Number of Significant Others		Pair by Pair Differences in Access to Significant Others		
	Controls	Experimentals	Higher than Control	Equal to Control	Lower than Control
Lower Achievement					
(- -)	3.5*	2.6	1/10	4/10	5/10
(= -)	3.2	2.6	2/14	5/14	7/14
(+ -)	3.2	3.6	4/8	2/8	2/8
			7/32	11/32	14/32
Equal Achievement					
(- =)	2.3	3.1	6/9	2/9	1/9
(= =)	1.2	2.6	6/9	2/9	1/9
(+ =)	2.6	3.4	6/8	0/8	2/8
			18/26	4/26	4/26
Higher Achievement					
(- +)	2.3	3.7	8/11	1/11	2/11
(= +)	2.5	4.2	16/27	5/27	6/27
(+ +)	3.2	3.0	10/25	3/25	12/25
			34/63	9/63	20/63
Total	2.9	3.4	59/121	24/121	38/121
(N)	(121)	(121)	$N=242, \bar{x} = 3.1$		

*Mean values are reported here even if there is considerable skew upward in number of citations.

relationship between pre-program resources and post-program achievement level by introducing information on significant others.

Table 5.2 reports both the mean values of the experimental and control groups as well as pair by pair differences in access to significant others within each resource/achievement condition. The most direct and telling comparisons involve the pair-by-pair differences. The mean values reported by program participation and by resource/achievement conditions, however, suggest some trends in the data. In general, the lower experimental achievers have lower access to significant others (14 are lower, 7 higher, and 11 equal). The only exception is in the case of high resource/low achievement experimentals, whose access to significant others slightly exceeds the level of access by the controls (4 higher, 2 equal, and 2 lower).

Exactly the opposite pattern holds for the high achievers. Those who became high achievers, beginning with lower resources (-+,=+), report the highest mean values in the study, (3.7 and 4.2 respectively). Conversely, the control group children report lower mean values (2.3 and 2.5 respectively). Among higher resource/higher achievement (+ +) experimentals, however, mean access to significant others is close to that of the controls (3.0 to 3.2 respectively).

For those experimentals who end up the same as the controls, the general tendency, still, is to exceed the mean level of access reported by the mothers of the control group children (18 out of 24 comparisons favor the experimental child.) In view of the random assignment to control and experimental groups, it seems reasonable to hypothesize that this post-program measure of significant others reflects an effect of the program. If it has not produced a measurable impact on our measures of academic achievement, itremains possible--and in our judgment likely--that it will produce other supportive effects for this group in the long run.

Table 5.3 examines in another way the relationship between levels of access to significant others and resource/achievement levels. Overall, 49% of the experimental children are higher in access to significant others than their matches. The incidence of such differences tends to be higher, however, in two out of three categories where the achievement level exceeds that of the control child and three out of three categories where the achievement levels are equal to each other, regardless of resource levels. Where the achievement levels are below that of the control child, in general, the incidence of higher access is extremely low in two categories and about expected in one.

Table 5.3

Percentage of Access to Significant Others Among Experimentals by Resource/Achievement Levels

Subsequent Standing on Index of Educational Achievement	Initial Level of Educational Advantage/Disadvantage of Experimental Child			Total
	Resources Below Control Child's	Resources Similar to Control Child's	Resources Above Control Child's	
	(Percentage of Higher Access Among Experimentals)			
Achievement Above that of Control Child	73%(11)	60%(27)	40%(25)	54%(63)
Achievement Similar to that of Control Child	67%(9)	67%(9)	75%(8)	70%(26)
Achievement Below that of Control Child	10%(10)	24%(14)	50%(8)	22%(32)
Percentage of Higher Access among Experimentals	50%	48%	39%	49%
(N)	(30)	(50)	(41)	(121)

The number of cases in each of the nine categories becomes quite small, but the pattern of findings is systematic in terms of direction of percentage shifts. In general, variation is more closely tied to terminal achievement than to initial resource levels, with the greatest access revealed by those who are equal in achievement to their pairs. These findings are consistent with those found be Sewell and others cited above.

Chapter Six: CONCLUSION

After five years of work on an experiment in the field, we sometimes long for the small-group laboratory and the twenty minute interaction. In another mood, we long for another five years in which to repeat the experiment reported here; for we believe that various departures from our ideal design could be reduced: We would specify more clearly in advance what we expected from the cooperating school systems; we would draw from a smaller range of schools in order to reduce problems of communication and transportation; we would recognize from the start the importance of the follow-up program.

It is probably inevitable that those who experiment in the field will be required to make adjustments to the wide variety of persons and situations that influence the interactions under study. We dealt with five school systems, with their different regulations, attitudes, tests, and methods of keeping records. We followed 195 pairs of students through several dynamic years, during which time national and local policies and the attitudes of many deprived families changed significantly. We have several sets of data on each child; but on none of the sets is our information complete. Although the rates of return on the original "pupil data form," on the continuing record of school performance, on the family interview, and on the mailed questionnaire have, in every instance been high (never under 80 percent), we have had to estimate some information and to set aside a few comparisons for lack of adequate measurement.

In spite of these difficulties, we believe that it is possible to speak with some confidence about various aspects of a "middle start" program. The most impressive lesson, to us, has been the demonstration of how important it is to pay attention to each new experience, each new challenge, if a child is going to learn to deal effectively with his school and life situation. It does little good to furnish him an opportunity for a stimulating and attractive summer program if he gets no support after that. Many of the students in our program did not know how to take advantage of opportunities in their schools and cities; they had little understanding of the financial aspects of college attendance, of necessary preparatory courses of study, of procedures for getting college catalogues and admissions papers. Without counselling, they would not have been able to put their motivation for further work into action. It is perhaps too much to say that critical learning situations during years 13 to 18 are, as we suggested in Chapter One, like links in a chain--if one of them is broken or missing, the chain does not hold. But this is not far from the truth. And we have learned that the same decisive continuity characterizes the experience of disadvantaged students once they get to college. Each year, new and often unanticipated problems arise because they are moving through a somewhat strange and unfamiliar culture. Any one of these experiences can break the continuity, bringing back a sense of defeat, or of hostility, or of isolation. These feelings can be of such strength that the desire to continue is sharply reduced or even crushed.

The ingredients for success in a middle start program, then, are these:

1. An exciting, new, stretching experience is required (in our case, taking them out of their home communities and onto a college campus), an experience that breaks the connection in their minds between school and failure, school and discrimination, school and boredom. During this beginning summer, new friendships are formed, many kinds of adult model are made available, a more extensive range of academic, artistic, and recreational activities than our participants have known before is laid before them.

2. A sponsor is needed, to show how new life possibilities are available, to explain each step along the way. In our case, the staff and director of the special opportunities program, college-age counsellors, and teachers and counsellors in the schools have served as such sponsors--although not with the continuous support that is needed. Neither budget nor our level of experience was sufficient to carry out this part of the program in an entirely adequate way.

3. A circle of "supporting others" is crucial. These are less likely to perform the instrumental tasks carried out by the sponsors. They serve, rather, to give emotional support, to recognize and encourage the changed motivations of a student, and, as he approaches college age, to show a readiness to help financially, if even in a small way. Parents are the most important members of this circle; but other adult relatives,

family friends, and adults in the community are also significant. Because this kind of support is essential, a program that deals only with the students will have a high rate of failure. We have sought, by means of reunions involving the families, by newsletters, and by interviews (which, although designed as a source of information, proved often to have a support function), to identify and strengthen the concern of the members of these circles. Parents have often shown an extraordinarily high level of support. Reunions have been well attended; some families have driven hundreds of miles, at a cost they could ill-afford, to demonstrate their concern; and the "testimonials" we have received document not simply appreciation for the program, but the strong motivation of most parents to encourage maximum educational effort, once a possible path for their child has been opened.

Each of these three ingredients is essential. When they are present, a significant proportion of children from disadvantaged backgrounds can have their skills, motivations, and academic performance records raised. Our participants (and our "controls"), it should be remembered, were drawn from disadvantaged families, but they were "visible" in their schools as children of some promise. One or more of three indicators pointed to them, not as strong students already destined for success, but as promising students who, with sufficient support, could move ahead. Their course grades were better than average, or their achievement-test scores were better than average, or their teachers saw them, despite poor grades and test scores, as persons with unused talent. We are unable to say that students who are "invisible" on any one of these indicators could be helped by the procedures we have described. It is our judgment, however, based on the performance of many students of "high risk" that there is a great deal of hidden potential among disadvantaged students, that carefully planned programs with continuous support through a number of years can be of great value to junior high school students who seem, by available measures, to be headed for failure.

Our total program costs were less than $2,000 per child, over a five year period. We did not get a strong follow-up program going for a year after the program started (and staff changes meant a weakening of the program in the last months). We did not have funds for a second summer, during which participants would have returned to campus--probably at the end of their sophomore year in high school--for further training. On a budget of $3,000 per child (scarcely the cost of one year in "prep" school for many persons in the upper and upper middle classes), we feel confident that half or more of participants in a strengthened program, similar to the one we have employed, would show significant improvement in academic motivation and performance. Total financial costs to society seen over a lifetime would be negligible. Indeed the "investment" would doubtless be returned many fold.

Our description of the effects of the program would be incomplete if we failed to mention several indirect consequences on persons other than the students themselves. We only call attention to these additional influences because we cannot document them with precision. Nevertheless, a flood of impressions, based on conversations, direct observation, unsolicited testimonials, and other evidences, make us confident that our summer and follow-up programs have significantly influenced a wide circle of persons:

1. Among our most active and important participants have been the college-age counsellors. They lived in the dormitories with the "middle starters," helped them academically, showed them how to deal with the new and sometimes confusing aspects of life away from home, joined them in recreation, welcomed them at reunions, and sometimes visited them in their high schools after their summer on the campus. We have no doubt that the college counsellors were an important ingredient in the program. But what we are emphasizing here is that they learned more than they taught. Almost without exception they wanted to return for another summer; their interest in race relations, in problems of the city, in education, and the like, was strengthened. Many of these counsellors were black students. Among them particularly there was increased power to see a difficulty, not only as

a personal problem, but as a public issue that could be dealt with only by understanding its structural sources.

2. The program has been scarcely less important for the teachers who have participated. They were selected on the basis of their interest and skill in supportive education; but many had had little direct involvement in such programs. They remark that their teaching is different; their sensitivity to the concerns of students increased; their interest in academic policy in their schools promoted.

3. Our evidence on the effects of the program for families of the participants is even less direct; but from our interviews and conversations at reunions, we see that parents and siblings (older as well as younger) have been drawn into the circle of influence. We have feared some negative effects at this point--alienation of a child from his family, antagonism of parents, jealousy of a brother or sister. Undoubtedly there have been such effects, but our techniques have not been sensitive enough to detect them; and we are confident that the positive effects are much the stronger.

4. Even more indirect, but worthy of mention, are the consequences of such a program on the institutions involved in them. In the course of a six year program, hundreds of school and college officials and teachers have heard about "special opportunities," have discussed it, and have helped to make decisions with respect to it. At least a score of Oberlin College professors and administrators have

had their minds and interests stretched. College-level work with students of disadvantaged backgrounds has increased and has been improved, partly because of lessons learned in the summer program. Administrators and teachers in the school systems from which our participants came have been led to ask themselves new questions about their own procedures and aims. Their support for the S.O.P. has contributed to an awareness of the range of problems in their schools and to support for other programs.

We do not know how to add up these indirect effects; and granted our values we may be inclined to exaggerate them. But we have seen their signs all around us. We have not seen signs of indirect negative effects. In the long run, such a program as we have described may have proved to be a "middle start" for all of us who have been involved in it, and not only for those who, in an inprecise way, we have called the "participants."

MIDDLE START

by

J. Milton Yinger

Kiyoshi Ikeda

Frank Laycock

APPENDIX

A I: A NOTE ON HUNT AND HARDT'S FINDINGS ON THE EFFECTS OF UPWARD BOUND PROGRAMS ON THE ACADEMIC ACHIEVEMENT OF NEGRO STUDENTS

Hunt and Hardt[1] report some general findings about the effects of Upward Bound (UB) programs on the attitudes, motivation, and academic achievement of Negro students. They find that there are no discernible differences in Grade Point Averages (GPA) between chosen (UB) and unchosen students (controls) subsequent to the program. They stress the importance of cumulative and innovative efforts to improve the academic achievement and survival of UB students in future programming and funding.

This note examines the research design employed by Hunt and Hardt to derive the findings and recommendations reported above. Insofar as major efforts to evaluate both types of Federally supported compensatory education programs--Upward Bound and Head Start--have relied on _ex post facto_ matching procedures to generate control group comparisons, a systematic discussion here may further reinforce the lesson which the social science community must learn. The lesson is that in future program evaluations, we must not rely on _ex post facto_ matching designs if we can encourage instead more powerful research designs that approximate true field experiments. Without such a standard we become embroiled in policy discussions which rest on findings which have weak foundations. Invalidating factors which detract from clear attribution to the educational programming itself abound in ex post facto studies[2]. Policy and funding questions are then based on blind guesses about next steps.

---

[1] David E. Hunt and Robert H. Hardt, "The Effect of Upward Bound Programs on the Attitudes, Motivation, and Academic Achievement of Negro Students," JOURNAL OF SOCIAL ISSUES 25 (Summer, 1969), pp. 117-129.

A major instance of such weak findings which have policy consequences is the Head Start program (See Circirelli, 1969; Smith and Bissell, 1970, and Campbell, 1970). The implication of the Westinghouse-**Ohio** study was that the program effect had been negligible and even harmful.

> In this context the Westinghouse-Ohio evaluation seemed to further demonstrate that public efforts to educate disadvantaged children were wasteful and futile. In February, President Nixon said that in view of the Report's findings, he would treat Head Start as an experimental rather than as an operational program.[3]

Those who are recipients of administrative largess may not, in fact, receive the maximum of direct and indirect benefits that the law intends. This is wasteful and futile. To assume that such consequences are intrinsic to educational intervention, however, is to run the risk that policy discussions and decisions will be based on very soft evidence. Systematic evaluation (based on research designs more powerful than ex post facto **matching** procedures) is both desirable and feasible for policy development.

---

[2] 2. Donald T. Campbell and Keith N. Clayton, "Avoiding Regression Effects in Panel Studies of Communication Impact," STUDIES IN COMMUNICATION, Department of Sociology, University of Chicago, 1961, No. 3 99-118, Bobbs-Merril Repring NS S-353.

[3] 3. Marshall S. Smith and Joan S. Bissel, "Report Analysis: The Impact of Head Start," HARVARD EDUCATIONAL REVIEW, 40 (February, 1970), 51-104. See also HARVARD EDUCATIONAL REVIEW, 40 (February, 1970), the whole issue on evaluation of compensatory education programs. These evaluations are derived from V. Circerell, et. al., THE IMPACT OF HEAD START: AN EVALUATION OF THE EFFECTS OF HEAD START ON CHILDREN's COGNITIVE AND AFFECTIVE DEVELOPMENT. Vol. 1. A report presented to the Office of Economic Opportunity pursuant to Contract B89-4536, June, 1969. Westinghouse Learning Corporation, Ohio University.

We would insist[4] with others[5] that the most powerful study design be employed.

Typical of ex post facto evaluation studies, Hunt and Hardt's research derive findings which are weak on at least two major points. The first is to rely on an ex post facto matching procedure for control group data. The second is the absence of repeated non-reactive measurements of academic performance involving the use of several (multiple) indicators. Instead, they rely on cross-sectional comparisons of two cycles of students at different ages.

The exact procedure employed for ex post facto matched controls and ten percent representative sample of cross-sectional data is described by Hunt and Hardt. All students in target programs (21 out of 215 in 1966) were studied. Effect of the program upon student attitude and motivation was measured by administering a battery of paper-and-pencil tests to students in the first week of the summer program, during the last week of the summer program, and during the spring of the following academic year. Effect of the program upon academic achievement was measured by collecting the student's Grade Point Average (GPA) from his high school in June before the summer program and in February after the program. GPA results were also collected for a control student not attending an Upward Bound Program but whose pre-program GPA was similar to the Upward Bound (UB) student.

---

[4] Kiyoshi Ikeda, J. Milton Yinger, and Frank Laycock, "Reforms as Experiments and Experiments as Reforms," upub. ms., Oberlin College, Oberlin, Ohio, February, 1970, 28 pp.

[5] D. T. Campbell, "Reforms as Experiments," AMERICAN PSYCHOLOGY, 24 (April, 1969), 409-429, and Richard O. Lempert, "Strategies in Research Design in Legal Impact Studies," SOCIOLOGY OF LAW 1 (1966), pp. 1-19

GPA data were collected by locating a grade recorder in each of 189 high schools during the academic year 1966-1967. Using the school records, each recorder first selected another non-UB student of the same sex and school grade whose GPA was identical or close to that of each UB student. Grade recorders were also urged to attempt to select a matched mate of the same ethnic or racial group and at the same income level. However, grade recorders were not always able to follow this suggestion. Therefore, the control group was quite comparable in initial GPA, but may not have been identical in regard to racial or social class criteria. Put another way, the 283 students who are the control group may not all have been Negro and may not have met the poverty criteria, but it is likely that most of them did.[6]

1. Ex Post Facto Matching and Regression Artifacts. In the Hunt and Hardt research, there is a strong likelihood that regression artifacts resulted from the post-treatment matching procedure. For one thing, the control subjects were less needy. This meant that the more needy were selected for the UB majority. If academic performance is positively correlated (even if weakly) with the family's financial status, then the control group would tend to have better grades, on

---

[6] Hunt and Hardt, op. cit., pp. 119, 127.

the average[7], in both the pre- and post-test situation. Choosing control subjects with family and academic profiles that really match those of the UB students would underestimate the potential and performance of the controls.

Campbell's discussion on the weaknesses found in the Westinghouse-Ohio design applies to Hunt and Hardt's procedure:

> (Post) Matching on several variables simultaneously has the same logic and bias (as is true of matching on a single variable). The use of multiple matching variables may reduce the regression artifact, but not remove it. It reduces it insofar as the multiple correlation of the several matching variables with the post-test is higher than the simple r of a single matching variable. Matching by means of qualitative dimensions or dichotomous variables has an equivalent bias. All such matching variables turn out to be imperfect indicators of the underlying variables we would like to match on. Parents' number of years of schooling have vastly different meanings from school to school, and within the same classrooms. Living in the same neighborhood or block means widely different things as far as the educational quality of the home is concerned...There is inevitably undermatching, in the sense that the population differences which one is trying to correct by matching are under corrected by the matching process...
>
> How can one tell which direction a matching bias will take? Only by having evidence on the nature of the population differences which matching had to overcome. Conceivably, in reporting a matching process, the researcher might neglect to say what kind of cases he found hard to get matches for, and what kind of cases existed in surplus in the control population...(in the Head Start study), it was the most disadvantaged Head Starters that were hard to match and that the controls were selected from generally more able populations.[8]

---

[7] William H. Sewell and Vimal P. Shah, "Socioeconomic Status, Intelligence, and the Attainment of Higher Education," SOCIOLOGY OF EDUCATION 40 (1967), pp. 1-23.

William H. Sewell and Vimal P. Shah, "Social Class, Parental Encouragement, and Educational Aspirations," AMERICAN JOURNAL OF SOCIOLOGY 73 (March, 1968), p. 559-572.

[8] Donald T. Campbell and Albert Erlebacher, "How Regression Artifacts in Quasi-Experimental Evaluations Can Mistakenly Make Compensatory Education Look Harmful," in J. Hellmuth, COMPENSATORY EDUCATION: A NATIONAL DEBATE. Vol. III of THE DISADVANTAGED CHILD, New York: Brunner/Mazel, 1970, ms., pp. 9-10.

In the UB study, regression artifacts operate in the selection of control subjects: the matches never quite seemed as disadvantaged as the chosen students for the UB program. If UB students also were chosen at the extreme of disadvantagedness, then regression effects of choosing at the extremes of a population also apply.[9]

The general effect of this approach to control groups is to obscure any significant gains on the part of the UB participants. This is due to the regression upward in the post-test of the control group in academic achievement. These regression effects can obscure the relative gains of the UB students. There are two ways to prevent this kind of interpretive impasse in the findings. The first is random assignment with or without pre-matching. With careful planning, Upward Bound programs can employ true field experimental designs involving random assignment of participants and non-participant controls. The fact that we have been able to develop a true field experimental design in evaluating a compensatory educational program which pre-dates Upward Bound efforts suggests that evaluation is possible, feasible, and even imperative. Hunt and Hardt did not have this option because, like Head Start, Upward Bound programmers did not pre-plan a systematic and hard-headed approach to program evaluation. Evaluation has been ad-hoc in both programs, leaving the program evaluators to adopt designs which are the most powerful under the given field conditions. We are not proposing that every program have a program evaluation component. Neither are we suggesting that systematic evaluation be the major basis for programming and funding of educational programs.

---

[9] Campbell and Clayton, op. cit., passim.

Campbell states our position most clearly and forcefully.

> The funds set aside for evaluation are funds taken away from treatment. The cost-benefit trade-off decision has already been made when quasi-experimental evaluation has been budgeted, or when funds are committed to any form of budgeting and accounting. Taking these evaluational funds, one could use nine-tenths of them for providing experimental expansions of compensatory instructions, one-tenth for measurement effects on the small experimental and control samples thus created... Since evaluation money would be used to expand treatment, the controls would not be deprived. In retrospect, we are sure that data from 400 children in such an experiment would be far more informative than 4,000 tested by the best of quasi-experiments to say nothing of an *ex post facto* study.[10]

Our own effort to evaluate an experimental compensatory education program involving approximately 400 students over a five year period suggests that efforts can be developed cooperatively among both those who administer and those who participate in a research study.

2. Measurement Errors and Biases in Cross-Sectional, Single Measure Studies. Lack of planning can result also in simplistic cross-sectional studies and on single indicators of an underlying variable. In the Hunt and Hardt study, academic performance is evaluated by Grade Point Averages for younger and older students in the program, with reliance on cross-sectional data for interpretation. It is not easy to rule out, with the UB study, peculiar historical effects and systematic maturational effects of going through high school. The finding that grade averages move in opposite directions

---

[10] Campbell and Erlebacher, COMPENSATORY EDUCATION, p. 25.

among White and Negro participants also remains to be interpreted. The likelihood is that both historical and maturational factors interact with race. Without systematic replication for several cohorts over a longer time, and without several other indicators of academic achievement--tests, rites of passage indicators of graduation from high school, admission to college, and survival in the educational and occupational structures--we remain uncertain about what Hunt and Hardt's findings mean on the ways in which these variables affect each other.

To be able to rule out biases and errors stemming from measurement and other design features, Blalock[11] suggests that multiple indicators be employed repeatedly.

> With a single measure of each variable, one can remain blissfully unaware of the possibility of measurement error, but in no sense will this make his inferences more valid. Though there is always the danger of becoming so hypersensitive to the possibility of measurement error that one becomes immobilized in the process, present practice seems to err in the opposite direction. Methodological studies... can help us see more clearly the nature of the step we must take if we are to become increasingly precise... I see no substitute for the use of multiple measures of our most important variables...[12]

Ex post facto studies typically neglect such considerations because of exigencies of time and funding.

---

[11] Hubert M. Blalock, Jr., "Estimating Measurement Error Using Multiple Indicators and Several Points in Time," AMERICAN SOCIOLOGICAL REVIEW 35 (February, 1970), pp. 101-111.

[12] Blalock, ibid., p. 111.

3

In addition, for cyclical programs repeating such measures (non-reactively, where possible) for a number of cohorts will uncover still better the invalidating influences stemming from both newness of the program and temporal influences.[13] Essentially, this means obtaining reliable, multiple indicators of change (or non-change) at several times. Thus, errors, from temporal effects, from maturation related to systematic fluctuation in indicators, to unique selection-treatment effects, could be evaluated. Compensatory education programs (or almost any educational programs) are suited admirably for repetitive longitudinal measures.

In our effort to evaluate a compensatory educational program for junior high school students we have sought to employ stringent design and measurement standards. We have multiple indicators of underlying variables related to household disadvantage (economic and educational) and of academic achievement. We have repeated measures (to a maximum of five years for a given cohort of students). We have three groups who have completed some parallel educational experiences. Again, evaluation which meets strict technical standards is not impossible in field studies.

In retrospect, we would like to extend the discussion by both Campbell and ourselves of what must be done to improve the quality of evaluation studies in improving education and related public services.

---

[13] Desmond E. Cook, THE IMPACT OF THE HAWTHORNE EFFECT IN EXPERIMENTAL DESIGNS IN EDUCATIONAL RESEARCH. (Columbus, Ohio State University, Final Report, Office of Education, Study Project #1757).

1. Program Evaluation Must Have Promising Programs for Assessment. We are not insisting that all programs be evaluated. We are insisting that selected programs which reflect crucial developments in policy be evaluated, where such evaluation is feasible and desirable.

2. Program Evaluation Must Be a Normally Funded Component. We are suggesting that programs reflecting critical policy changes be evaluated as normal practice.

3. Program Evaluation Must Employ the Most Powerful Designs, Field Conditions Permitting. This usually will mean careful pre-planning for field studies involving true experiments with long-term measurement.

4. Ex Post Facto Matching Procedures Must Not Be Employed Where the Three Criteria Above 1-3 are Favorable.

We are not calling for a moratorium on ex post facto studies in education. We are asking that premature interpretations from ex post facto studies be laid aside in favor of systematic support of more stringent evaluation studies.

We favor studies which speak to critical policy problems and which use stringent hard-headed evaluation. Where such critical studies reflect some deficiencies in planning and execution, such evaluations should be strengthened wherever possible. This would mean laying aside funds for re-evaluation, as necessary, and for program and evaluation shifts in response to incoming information.

In compensatory educational programs, such as Upward Bound and Street Academies, we suggest that a wider range of programs be evaluated. These programs should reflect systematically varied approaches to deducing talent loss. They should work with the very educational structures and populations who seek to find effective solutions to the problem of loss of talent and leadership. One program, instead of focusing on counseling and remedial or enrichment efforts in the summer and follow-up during the academic year, simply can lay aside funds for going to college for given students.This is a variant of the GI Bill. No child will then be denied an opportunity to go to some college for lack of money. A child and his family can be notified of this financial opportunity. Subsequently, the child's performance can be evaluated.

Another program can focus on counseling, as in Talent Search. Still another can follow Upward Bound. Still another can use Street Academies. Combined residential/non-residential, remedial/enrichment, counseling/non-counseling programs can develop in a number of institutions of higher learning and public schools. Cooperative programs can be evaluated for relative richness in programming and funding. Some will provide little but fiscal help at the end of high school. Others will add to that assistance. Still others will expend heavily in professional services to prepare the student for college. Still others will focus only on one or two services. Still others will involve shoe-string operations with a great deal of information transmitted by individuals who come out of the same social and cultural experiences.

The only research design stipulations would be (1) the initial talent pool must come from a common source--a major metropolitan or rural area; and (2) some common yardsticks of academic performance and achievement must be employed within a longitudinal study. Different public schools or public school districts can coalesce with institutions of higher learning to work on one critical feature or another. Where institutions and schools compete, variants may be encouraged which are comparable except in one critical policy feature. In this way, random assignment can be employed to assign students from a common pool into cooperative programs with some potential for showing which policy feature or features have the most impact. In this way, no child who needs assistance would be denied some aid; given our desperate concern with talent loss, no approach which appears to have merit need be denied a standing chance at evaluation; and, evaluation can proceed in the most hard-headed way desired by both policy-maker and social scientist-evaluator.

We thus hope that compensatory education programs become "more experimental", that they reflect the deep convictions of those who seek to improve educational opportunities for all of our children by allowing program variations which permit a fair comparison. At the same time, we can keep to strict standards of program evaluation so that our findings can speak more directly to policy issues. It is in this spirit that we have examined Hunt and Hardt's evaluation of the Upward Bound Program.

A II: CRUCIAL DECISIONS IN PREPARING TALENTED URBAN YOUTHS FOR COLLEGE

What educational decisions face an educational institution today when it seeks to provide community services in urban areas? Can such institutions take up the challenge of the communities and populations around them, which are subject to urban blight and limited opportunities? Six years ago, Oberlin College started a twin-pronged effort to provide special educational opportunities to talented youths who would not normally enter college or earn degrees because of limiting social and economic forces. One part of the program has dealt with junior high school students, the other with college students. Our major goal has been to build a larger base of effective local leadership. We hope that after having gone to college and into appropriate vocations afterward, these youths will be able to contribute more richly to their own groups and communities and to society as a whole. This article discusses the decisions we made as we developed our pre-college program of encouragement and support. (Our experiences in the college-level program will be reported elsewhere.)

It is well established that talented youths in urban public schools limit their plans and preparations for college because of inadequate financial and educational resources in their families and neighborhoods. The most severe limits are placed on minority youths in economically impoverished areas. Without some significant educational and financial efforts, most of these youths will not become leaders or make effective contributions to society. When Oberlin began its program, we had little information about how institutions might break through the limits imposed by residential, social, and economic patterns. We began with the best we knew, and we have been immersed ever since in a cooperative program with the public schools and the participating families. This cooperative program is teaching us many lessons. This article will try to point out, with the wisdom of hindsight, the crucial decisions that demanded attention and that have dictated the results that we are now analyzing.

First, briefly, the facts. In 1963, officials of the Rockefeller Foundation asked Oberlin (and Dartmouth, Princeton, Antioch, and others) to try to help poor but bright students get to college--students whose backgrounds rarely nurtured the dream of higher education or, indeed, of even completing high school. The Foundation was hoping that something could be tried on college campuses to enable promising students to reach their full potential. The efforts of the Rockefeller and other foundations contributed to the development of large-scale federal assistance for "Upward Bound" programs. In 1963, however, the whole proposal was novel and exciting; its outcomes were quite uncertain. The Foundation did allow discretion, so that each college could pursue its own pilot efforts on the basis of its own judgment. The hope in all of the programs was that exposure to properly designed pre-college experiences would lead to ready admission and able performance in high-quality liberal arts colleges and universities.

Oberlin chose to bring a group of seventh and eighth graders, about two-thirds of whom were boys and two-thirds Negroes, to live for six weeks on campus during the summer. They were encouraged to study hard, to play hard, and to receive as much stimulation and counseling and encouragement as the College could command. Since that first summer we have decided to follow these boys and girls more closely than we had planned in the initial phases of the pilot program. We now employ various supports: counseling, extra academic work for both enrichment and remediation (through tutoring, placement in other summer programs, assignment to private schools), family contacts (through parent clubs and newsletters), book services, visits, reunions, and college placement advice--all to keep alive the spirit of the first summer. We have continued our summer curriculum for different groups of seventh graders through six years, and we offered the seventh one in 1970. The pupils have come from two large cities (Cleveland and St. Louis), two smaller cities near the campus (Elyria and Lorain, Ohio) and Oberlin. Each year the summer staff for our 65 to 70 pupils has included more than twenty teachers, college counselors, and other adults. Lately some of the former participants, including those now in college, have come back to serve as junior counselors, aiding the adult staff in working with the younger children.

To assess the impact of this program, we are studying intensively the first three groups of seventh-graders who have had this experience, 194 pupils in all. Oberlin wished not only to help these particular children, but to find out whether our form of help would enable other such youths to continue on into higher education. We wanted to be able to assess our experience in such a way that other colleges and public schools could see which techniques had worked and which failed for us, and which ones would probably work for them and which ones not. We wanted to assess the full process--from approaching school districts and locating the pupils, to the end of the follow-up. (A grant from the federal government is supporting this research dimension of the project. The research involves not only the pupils who attended the first three summer programs, but also an equivalent, matched group of pupils who did not. We shall report on this research elsewhere.)

Now that we look back, it is clear that our program could have been different. Its form was arbitrarily set by decisions that followed in succession as the plans were being developed, staff selected, and later as the operation moved along. Some decisions came thoughtfully and systematically, some were hurried or compromised because of deadlines or emergencies. It is to examine the most important of these decisions that is the burden of this article.

1) What could an affluent, liberal arts, highly academic college do to help talented children from less affluent families?

For Oberlin this question was easier to answer than it might be for some other colleges. Oberlin was the first to admit Negroes to higher education. Since well before the Civil War a small but steady number of students had come from minority backgrounds to the campus. The Rockefeller support merely required that the College augment its historic commitment. Various members of the faculty, particularly those who had worked with young Negroes, pointed out that it would not be easy in the 1960's to win trust or to inspire effort among children who live in the urban ghettoes. Higher education may be valued in general, but harsh reality has often dictated to such youths that they must lower their aspirations and their actual plans. Such adolescents would be wary of tricks and sensitive to condescension. Their families, caught in the cycle of poverty, might be suspicious or apathetic. (Later experience has shown these fears to be unfounded.) Oberlin has obtained maximum help from specialists on its staff who know the poor from first-hand study, from students and others whose homes were in slums, and from leaders in the current campaigns in behalf of Negroes. Otherwise, the cultural and social chasm that separates those who live in middle-class life styles from those whose life styles are less affluent could not have been bridged. We took as a premise the belief that economic impoverishment should not be the basis for judging the worth of those who live in poverty. Anything that approaches a good-hearted but casual philanthropy, which lifts the hearts of the givers but treats the poor as wards of their more prosperous fellows, is almost certainly inadequate today. Moreover, anything that would limit a Negro's pride in his background would bring sharp retaliation.

We have attempted not to get caught on either horn of this dilemma. We believe that the children and families should be proud of their deep social and cultural "roots" and traditions. Curriculum, staffing, and follow-up programming have reflected this view. On the other hand, we have sought not to understress the demand for upward educational and occupational mobility on the part of the children and their families, even though mobility may entail personal and fiscal costs. The hope was, indeed, that these children and their families could serve as significant models for others with talented children to invest actively in the future education of their children in effective ways. This would mean the acquistion of personal and social skills to begin to understand their own family situation and of others like themselves and to begin the long-term process of navigating a social and educational world which has its own demands and values. These demands and values may or may not support past values and experiences. Personal and interpersonal strain is built into social mobility. The challenge in summer and follow-up programming has been to develop modes of mutual acculturation so that assimilation would not mean loss of past self and family identities, as much as a healthy mix and blend.

Besides the question of whether or not the experiment did create a change in the academic lives of these children, there is an essential question in program evaluation involving compensatory education. The underlying question is whether these children, plucked by chance from social and cultural traditions which may be at variance from those values and practices of institutions of higher learning, can and have come to terms with both their old and new academic/cultural experiences. Coming to terms means absorbing and integrating their new intervening experiences into the ongoing demands of family, peers, and school/neighborhood arrangments with which these children must cope. Our sketchy analysis here represents the beginning of detailed analyses of that personal effort by the child and his or her family.

Through these years we have also wrestled with this fundamental issue: how to insure that the minority communities themselves will be able to promote competence and encourage responsibility as they try to meet the demands for upward mobility. From the beginning we stressed effective community leadership as the final goal of training for college and beyond. We tried, therefore, not to wrench these youths from their home settings. Instead we tried to help them shuttle between divergent social worlds, hoping that they would develop the capacity to cope with both traditions. We are not yet finished with our debate over how to promote integration into the mainstream of American life, how to maintain among our participants identification with the old background while they participate pridefully in the new. This issue troubles all of us. Yet our policy has been clear that youths must confront both their local communities and the larger world.

2) How radically should a college intervene?

By any standard, our program has remained in the middle ground of college and public school policy. We have not tried revolutionary methods nor overturned established patterns. But neither have we been content to follow different groups of pupils through their regular experiences, measuring how their schools and homes worked together and watching to see if they ever got to college. This latter course might have added something to our understanding of the 'disadvantaged', but it was not what the Rockefeller staff or our faculty wanted. All sought to increase the supply of talented youths from minority backgrounds. We decided, therefore, to explore what we thought were some reasons why so few pupils from the ghetto go to college. If we could identify critical reasons, we might begin to answer the question whether talent loss as a result of the established pattern of schooling might be reduced.

Because other institutions were selecting pupils who already were on their way to higher education, picking high school students in their last year or so of an approved college-preparatory curriculum, Oberlin decided to dip lower down into the tested talent pool. We included pupils whom both the public schools and selective colleges would put onto the non-college tracks. We wanted to see whether selective tracking in schools and admission practices in selective colleges were missing talented students who could make it with some special preparation.

We urged the schools to look more carefully for talented pupils in their classes, and to encourage them to prepare for higher education. There is no alternative to direct school participation in this process if we are to insure proper identification and support on a routine, permanent basis. Public school staffs must be able to find and stimulate these pupils. Ad hoc procedures and special crash programs from the outside cannot easily be incorporated into regular routines or transmitted across the legion of districts facing this baffling problem. Our effort to reach deeper into the talent pool was novel for its time, and was a strategy that could increase the number of minority pupils attending college well beyond that which would respond to less aggressive efforts.

We depended upon techniques and programs that most school districts could adopt without major modification. They might not try the total program, but they could adopt parts of it to test within their local setting. Smaller classrooms, individualized instruction, active training toward self-reliance in both academic and personal affairs, self-determination of school and home rules--all of which were important aspects of the summer programs--are no longer novel. But they were excellent opportunities to challenge a staff which had been selected from local school districts in the cooperating neighborhoods. We also wanted to try an experiment the results of which could be used by other colleges working with fairly traditional schools. We did not thereby ignore or oppose more radical innovation or intervention; we chose a middle ground policy because it seemed to hold the most promise of aiding the large majority of public schools serving the children of the poor. Today we remain convinced that this whole question is one of the most important that any college must confront as it thinks about providing effective community service. Fssentially, we have sought to determine which kinds of supplemental assistance, provided under a cooperative program, are effective and can readily become normal procedure to help any eligible child in the public schools.

3) Which goals should the college encourage?

The Rockefeller Foundation staff believed that the primary effort should be to demonstrate that minority youths with talents and preparation matching those of majority youths could succeed in highly demanding colleges. Because we decided to work with a group of students at the fringes of established requirements for entering academically demanding colleges, however, our major aim was that these students should enter a college that matched their talent and interest. We also found that the traditional college-bound curriculum was not always the one needed. Preparation to enter trade schools, technical institutes, and community colleges was also valuable, if the basic goal of extending schooling beyond twelfth grade were to be reached. So we moved to broaden our stated aim to include graduation from high school by whatever program would encourage further study or training. For some pupils we faced an even more basic problem: to see that they stayed in high school long enough to graduate. In general, however, it happens that the majority of our pupils have so far chosen a four-year college for their goal.

4) What age pupils should the college choose to work with?

When Oberlin faced this question, no college we knew of had dipped below the eleventh or twelfth grade. Dartmouth and Princeton, for example, used their Rockefeller grants to bridge the gap between secondary school and college. Since we wanted to change attitudes as well as skills, we believed that a summer program with follow-up afterward should make more difference to a younger pupil, whose attitudes were less likely to be fixed. Certainly if we waited until eleventh grade, some talented pupils would already have dropped out of school entirely, or entered non-academic terminal programs. Public school people agreed with us that we should be able to prevent many drop-outs if we chose to work with seventh or eighth graders. So we decided upon the summer between seventh and eighth grades. Younger would have meant trouble during the summer away from home at a strange campus; older would lessen the chance to influence a pupil or his family or his school. We were particularly eager to reach him before he had been 'tracked' at school into a non-college program or into one that would make some other schooling past graduation unlikely. We found that in some of our schools tracking had already started as early as seventh grade. But in most schools it was not until senior high school that serious tracking started. In the years since we began our study, of course, intervention has reached even into the pre-school years, in "Head Start" and other programs. Colleges may decide to work with younger and younger pupils, for the same reason that Oberlin chose twelve-year-olds.

5) Which kinds of pupils should be chosen?

We deliberately decided to choose pupils whose chances of preparing for college were small. We had to specify exactly which types of pupils we wanted to find out about, and then we had to figure out how to choose a small number of particular children to represent these types adequately--the problem of sampling. First of all, we established criteria to look for children from homes within a specified family income range. We deliberately picked more boys than girls (2 to 1), more Negroes than whites (2 to 1), and we avoided all but a small percentage (less than 10%) of students with behavior problems. We concentrated on boys because we wanted to strengthen the male image that is often weak in ghettoes, and because we knew that it is boys in particular who have to prepare to make a living. We concentrated on Negro males because they were needed the most for community leadership. Historically, Negro males have been the ones most systematically denied opportunity. Many of the non-Negro pupils belonged to other minority groups (e.g., Puerto Rican, Appalachian whites, children of Eastern European immigrants). We found it harder to decide about pupils with severe behavior problems. They need unusually sharp attention and highly skilled helpers. Without close support, they can easily disrupt the best of programs. We declined to leave them out entirely, but we limited their number severely, lest we confound enrichment and therapy.

A more technical aspect of selection worked better than we had expected. We wanted to put more authority into our eventual conclusion. This meant, ideally, picking pairs of identical pupils and then assigning one pupil at random to our enrichment and the other pupil to a control group for comparison. We didn't, of course, find identical pairs. But we did set up a careful list of traits to try to match, such as family structure, parents' occupational status, and previous school performance. We presumed it would be easy to assign pupils at random but difficult to avoid hurt feelings among parents whose children were not picked for the enrichment program. We were pleased at the extraordinary cooperation of the schools. We had asked for a large slate of nominees from school guidance counselors and teachers, because we feared that the schools might nominate only their model pupils. In most cases we did get a sufficiently large slate to permit good matches. Final selection of the pupils and their matches was made by Oberlin College staff, not the schools. We are convinced that random assignment is worth the effort required, because it permits a cleaner research design for the experimental and control groups.

6) What should the concentrated intervention include?

The pupils chosen dictate some of what can be done, and the goals of a program suggest more. We asked for bright pupils who had college potential, but whose backgrounds made college very unlikely. For goals, some of us argued that the greatest need would be to work intensively on the basic academic skills of language and number, by giving heavy courses in reading and writing and in mathematics. Some others of us wanted to plug cultural gaps by filling the summer days with art, music, drama, and other activities little known to the pupils. Still others of us pushed for a counseling emphasis in order to focus on career skills and attitudes. There is also the very practical question of how successfully a pupil could integrate new values into the home environment he would return to after the summer at Oberlin.

In the end, despite the large staff and the complex schedule required, we combined all three. Each child took at least one basic 'solid' course; there were other elective courses; the weeks and especially the week-ends were full of field trips to museums, concerts, and plays on or off campus; trained counselors worked individually with each pupil. In the second and third summers the staff changed its approach as experience suggested, without ever settling on any one exclusive emphasis. For example, the courses required of each pupil were shifted, the time available for personal counseling was increased, the extent and nature of field trips were varied.

7) How much follow-up is needed, what kind should it be, and for how long must it continue?

It was clear that extensive work would be required to keep the enthusiasm the summer had produced. We feared a let-down, since we were sure that six weeks at age twelve would not change the life of a person who went right back to old haunts. In an effort fo find programs that would prevent such a let-down, we canvassed the literature on intervention, we queried experts, we thought out our own devices. For example, we picked one or more counselors and teachers in each school to which our pupils went, to serve as liaison with Oberlin. We were provided with all the school data we needed--though not always automatically--on our pupils and the control pupils. For follow-up we arranged for counseling, tutoring if needed, field trips, visits back to the campus. We sent trained visitors to meet the families and gradually tried to open up the whole question of college with them: opportunities, financial aid, prerequisite courses needed in high school, and so on. Some schools had such encouragement already built into their regular programs (notably the Banneker district of St. Louis). We planned much briefer workshop reunions for both children and parents at Oberlin, trying to center on college requirements and intensive counseling. It was our guiding assumption that the summer could only make a lasting difference if the home and school tried to carry on the momentum afterward. We are convinced that no college should attempt a summer program unless it is realy to extend its help well beyond the initial, concentrated effort.

The perfect situation for a college would be to have data from experts already at hand, to point out ahead of time which combination of home and school and community supports would be effective. It turns out that such information is pitifully meager. If we succeed in our research, we hope to have something concrete to recommend. Meanwhile we concluded that we would have to supplement the ideas that research had already validated with a variety of hypotheses derived from logic and common sense. By urban, public school standards, we also chose more intervention than might prove to have been necessary, and yet, by preparatory school standards, the effort is weak. Our dilemma was a common one where the theoretical basis for such an enrichment program is not sufficiently developed, where one cannot wait to act until all desirable preliminary information is available.

8) What records should be kept?

Because we worked with five different school districts, and within them in many schools, the variety of tests, reports, and other information in our files has become voluminous. There are standard tests of intelligence and achievement, protocols of long interviews, family data, teacher comments, summaries of school activities, some biographical statements, even extensive lists of address changes. Because the control children were in the same districts, some methods of obtaining data were not open to us. We could not do anything for our pupils that would affect the control pupils too. Therefore, we did not undertake, for example, a single test for all our pupils, even though this would have simplified standardization for us. Nor did the schools alter their routines drastically, by doing something to all the pupils, merely to provide us with more data about our own. Here more than anywhere else the differing demands of follow-up (do everything possible to help) and research (keep the line between experimental and control conditions clear) caused concern and occasionally dissension.

Where necessary, we have chosen to protect the research design. This is not the only choice open, especially where research is not the prime consideration. But research sometimes requires that for the moment we restrict the urge to help, even when some school officials or parents do not thoroughly understand or agree. We hope that after we evaluate our experiment, some of the tests or other devices we employed may be introduced routinely into these schools, provided we demonstrate their worth.

9) To what non-academic areas should the college direct its attention?

This question carries forward the more general questions about follow-up. Because of the structure of many Negro families, we devoted not only follow-up attention but research activity to family relationships. The family is the earliest social experience, it can offer powerful models, and it attitude toward education is crucial. A long interview with a parent (usually the mother) and other inquiries into home life have become critical parts of our research. Just as logically, the home has been a center for encouragement and support during follow-up. The gap between the middle-class college and ghetto home life can be very wide. Many of the interviewers whom we trained for our research were drawn from the areas around the schools; and we depended heavily in the program on persons in the schools who already had the family's confidence. We were fortunate, after the project was under way, to have for the college summer staff many persons who had first-hand experience with ghetto living. It is a rare professor who has this direct experience. We could call upon our regular faculty for the expertise that comes from knowing the literature and from participating in research and academic study of minority problems. But for some tasks, especially those that must be done in the ghetto itself, it is strategic to turn to persons who know it intimately and who have been accepted by its residents.

10) How can the college continue to offer help when conditions are changing?

It is clear that colleges stand at something of an impasse now as they seek to provide substantial community service. There is frequently a desire to add to the traditional volunteer work by students and staff with some local community groups and agencies, payments in lieu of taxes to increase the level and quality of community services, and efforts to build educational programs with public schools, among other services. But in spite of this desire, colleges may feel unwelcome, now that there are greater community efforts at self-determination in local affairs both in the schools and in the general governance of the community. Negro pupils and Negro parents may appear to want more in the way of relevant studies in Afro-American history and literature; more Negro staff may now be available to fill the gap in supporting the educational efforts of a population long subject to discrimination and prejudice. Public schools themselves are now undertaking more programs, often with federal support, to supplement the missing elements in previous educational work. Intervention and aid by and with colleges and universities may appear to be less needed, or perhaps needed in different ways.

On the other hand, our experience convinces us that colleges can develop the kind of cooperative technical assistance that will enable school staffs to cope with pressing educational problems and to enable parents and their children to take care of educational matters in an informed and skillful manner. Several clear educational policy gains can be tentatively identified: (1) increasing ability of school staffs to search for talent for college in the public schools we have worked with; (2) increasing recognition by elementary and secondary schools staffs that they need to coordinate their work more closely to aid individual students; (3) training benefits accruing to the public school personnel who participate in the summer programs and follow-up activities; (4) training benefits for both undergraduates and graduate students in education who begin to understand the complex ties between programming, school staffing, and family participation; (5) further educational benefits to older students as they become 'big brothers' and 'big sisters' during the summer counseling.

Thus, a cooperative educational program is evolving. Our hope is that from it will emerge a better understanding of the educational processes that influence the fate and future of talented urban youths.

A III: REFORMS AS EXPERIMENTS AND EXPERIMENTS AS REFORMS

Research Design. A brief discussion is reported here of the treatment variable(s) and the associated field experimental arrangements made to evaluate the cooperative educational program.

a. The Treatment. The cooperative educational program can be described as a long-term educational project that attempts to meet the educational needs of talented youths from economically impoverished settings. Our first contact with participants is at the end of their seventh grade (a grade level significantly lower than in most programs of identification and support for talented, but economically less advantaged youths). The educational program (called SOP by participants) begins with an intensive but pleasant summer's experience on a college campus. This is followed by structured (and informal counseling when requested), newsletters, reunion workshops, subsequent summer follow-up in academic remediation, enrichment and staged educational-occupational career planning, and by related "follow-up" activities through the next several years up through the first year in college. This range of supportive interventions is expected to produce significant effects in academic achievement, staying power in school, higher educational plans and attendance, and occupational plans.

On six successive summers, with both external funds and internal college funds, we have brought about 65 seventh grade children (381 in all) to the campus for six weeks of study, play, counseling, and various other activities which stretch well beyond their usual range of experience. A staff of twenty-two persons, including twelve teachers and ten college level residential counselors (now including former SOP students in colleges and universities) are involved in providing an intensive and novel summer experience. For five summers, this residential component has been the keystone of the program. For the sixth summer, and probably for this summer, a nonresidential program is involved. The summer staff (drawn from the participating schools) also take part where they can in long-term follow-up on a formal and informal basis with other school staff and parents.

b. The Design. Our cooperative research effort involves a study of the impact of the SOP program on the educational careers of the 194 students invited to participate in the first three years of the program. Our final research design is a patchy institutional cycle design (Design #6 or the Fisher After-Only Design) with precision matching before random assignment. The method of precision-matching is described in another paper.[1] As Campbell and Stanley suggest, a true experimental design (A Fisher After-Only Design) is more likely to control invalidity stemming from history, maturation, instrumentation, regression, and mortality, among internal sources of invalidity. Except for long-term fatigue effects on researchers as observers (as instrumentation), internally invalidating effects appears to be well-controlled. In general, we have been able to obtain consent to employ the most powerful assignment procedures to generate equivalent comparison groups.

---

[1] J. Milton Yinger, Kiyoshi Ikeda, and Frank Laycock, "Treating Matching as a Variable in a Sociological Experiment," AMERICAN SOCIOLOGICAL REVIEW 32 (October, 1967), pp. 801-812.

The method of pre-matching before random assignment limits our ability to generalize to the total nominated and eligible list of students. Matches cannot be found for all the students who are nominated. We think our initial obligation is to establish the size and direction of differences within the study itself. Later, as programs of talent search become better able to scan the range of students more effectively, generalizable studies could be undertaken on a larger scale.

We recognized that there would be "reactive leakage" between the chosen and unchosen children in the schools selected to nominate lists of eligible children. The students nominated for the program are from similar ability-track levels, share some classes together, and thus may know each other within the school and neighborhood setting. We developed a patch of "nonequivalent control group" of post-matched students from schools post-matched with our chosen schools. Campbell and Stanley label this design as Design #10, The Non-Equivalent Control Group. (This patching would not be necessary if we had pre-matched impacted schools before pre-matching students within and across schools.)

Chart I outlines the typical cycle of nomination, selection, participation, and evaluation of the participation (or non-participation) in all three cycles of students under study. Within a given cycle, Cohort A includes both the chosen (Group A) and unchosen students (Group B) who were assigned from the chosen school invited to nominate children for the program. Cohort B includes the post-matched students from the post-matched schools (Group 1B) as well as a set of post-matched students in the chosen schools (Group 1A). The latter were sought because no student who was nominated provided an adequate match in the pre-matching effort. Both cohorts are followed over time through archival records of academic performance and by interviews and paper-and-pencil inquiries of the heads of households of the children in the study. The archival checks are less reactive; the other checks may be quite reactive in influencing study results.[2]

II. Obtaining Consent and Cooperation. The essential problem in mounting a hard-headed evaluation study through external research is that of inter-organizational coordination. With political and personal vulnerability a constant threat in either intra-mural or extra-mural research, public school staff, for example, must be assured that any research does not harm their personal and official integrity and security. To open intra-agency operations to detailed scrutiny, especially on ability to "deliver the services" means that the agency staff must have or must develop a political justification which says that "honesty is the best policy." Researchers must be responsible in providing such justifications. In the absence of such justification, such staff are unlikely to share organizational knowledge or secrets of staffing and operations.

What researchers actually do to agency operations is to increase the organizational work of the staff by demanding performance out of the ordinary routine. To sanction empirical inquiry

---

[2] Donald T. Campbell, "Reforms as Experiments," AMERICAN PSYCHOLOGIST 24 (April, 1969), p. 409-429, especially p. 410 with regard to ways of checking on plausible rival hypotheses.

A PATCHY INSTITUTIONAL CYCLE DESIGN (MODIFIED DESIGN #15)

Experimental and Control Groups within Cycle (repeat for 3 cycles)		Experimental Treatment (multiple treatments extended over five years)	Measurement Non-Reactive Measures	 Reactive Measures
pre-matched* Cohort A (students nominated, & assigned from chosen school)			Design #6: After-Only Fisher Design expanded into Design #15	
Group A	Selected randomly to participate	Receives treatment		
Group B	Selected randomly to serve as controls	Does not receive treatment	Archival measures from school records	Interviews and paper and pencil measures
Cohort B (students post-matched to Group 1 students)			Design #10: Non-reactive, post-matched, nonequivalent control group expanded into Design #15	
Group 1a	Post-matched within chosen school			
Group 1b	Post-matched from unchosen school			

---

* The first cycle of students in 1964 were assigned by simple random assignment rather than by pre-matching before assignment. Post-matching of nonequivalent matches within the nominated, eligible group has also been made.

regarding staff operations means that accountable staff must be sure that political vulnerability must be minimized. To sanction such study also means requiring counterpart agency staff to work closely with researchers to complete newly assigned functions which the staff may not have had to perform at any time in the past. Such extraordinary staff assignments require a clarification and crystallization of reciprocal rights and duties[3] involving researchers, responsible supervisory staff, and line personnel (above the level of clerical staff).

This three-way interaction requires that researchers take great care in coordinating consent with central supervisory staff and cooperation with line staff. Any implication of ad-hominem research places abrasive strains in the three-way interaction. Managing this problem places high role strain on staff involved. We have found this to be so in our initial and early stages of the research. Moreover, any additional costs in time and personnel required to service the research (or even the programming) component become the basis for intensive discussion as to reciprocal rights and duties.

Eisenstaedt's general discussion of institutional change and the specific processes involving the legitimation of new goals and roles within on-going structures aids us in our analysis. He suggests that two major processes must be coordinated in

---

3. "A right is a legitimate expectation entertained by a person in one position with respect to the behavior of a person in another position. From this point-of-view of the other person, this claim represents an obligation. 'Right' and 'obligation', therefore, are simply different definitions of the same relationship." Kingley Davis, HUMAN SOCIETY (New York: Macmillan, 1949), p. 87.

"reforming" internal organizational work to conform with external research demands. The first is to obtain informed and supportive consent through status-equivalent contacts. In a nominally hierarchical structure, this means beginning with that level of top staff who can consent to cooperation without review of others in that decision. The second is to clarify and to crystallize role demands and goals and beliefs which justify such new demands in ways supportive of the primary mission of the staff. Such "crystallization" of goals, beliefs, and roles must "answer their (staff) need for some general stability or more specific needs and problems."[4]

At the specific level, a decision to consent and cooperate in detailed scrutiny of staff operations must take into account the willingness of the agency staff to risk their careers and their agency functions to the possibility of negative evaluations and recommended changes in staffing and administration. At the general level, the primary aim and sub-missions of the agency is being evaluated with the possibility that such missions may face redefinition. This evaluation may come down hard on the livelihood and official functions of those directly involved in allowing research access to agency operations. One might surmise that this means that agency staff will never "stick their own necks out" in intra-mural or extra-agency research, if it is their own necks. Retrospectively we do not believe this is the true state of affairs.

---

4. S. N. Eisenstadt, ESSAYS IN COMPARATIVE INSTITUTIONS (New York: Wiley, 1965), p. 30. See his full discussion on pp. 3-31.

The more accurate proposition is that agency staff will not "trade off" on evaluative, policy research unless they can see a significant improvement in agency services in which they are involved directly and personally. If research is ad-hominem in a direct fashion, it is the loss of their jobs and good names. If research may affect their livelihood and functions, but they can see a chance for significant improvement in services and outcomes, and they personally can aid in bringing it about, staff will be willing to undertake the time and effort (costs) of justifying the evaluative research effort. We approached agency staff with a full recognition that agency staff security could be involved, but we fully discussed the primary mission of the staff in relation to present outcomes and future outcomes. In "impacted schools," both central staff and local school staff have been desperate and are desperate now as to how they can meet their primary task of educating the young. It was this desperation and pessimism which served as a backdrop against which we have discussed and even argued with the staff about hard research on alternative programs. The staff have consented and acceded to reforming their programming and evaluative efforts in the cooperative educational program with the college, even at the cost of personal pride and official retractions. We also have stubbed our toes in coming to know the organizational processes. We have done our share of eating personal and official crow in both programming and research coordination. All in all, administrative and teaching staff have consented, despite constant feelings of personal and official threat because they do wish to service the young fully and effectively. Even if it has hurt, both in time, effort, and organizational and personal pride, they have given more than we have ever expected so that the research could be completed. They have been willing to trade off livelihoods in the hope that something positive can come out of systematic programming and evaluation.

In general, unless researchers can clear the channels through status equivalent arrangements and provide justifications which promise some satisfactory trade off of gains over costs of work involved in research, agency staff will not consent and cooperate. The remainder of this paper describes work involved in clearing the channels for consent and in providing justifications based on cumulative knowledge about what the staff must cope with in adopting new role demands of systematic research.

a. Clearing Through Channels. We learned very quickly that what we thought was normal and ready clearance through channels was insufficient in obtaining consent and cooperation from staff lower down in the staff hierarchy. We have had to loop upwards within the hierarchy a number of times to insure programming and evaluation efforts at the local school level. For one, we have found school hierarchies to be more "horizontal" than regulated directly from above. Local staff always have local discretion in determining the justice and legitimacy of an external request, even with central staff approval. In matters involving who does have the final say, local school staff can combine to overpower central staff efforts to obtain consent and cooperation for external representatives like ourselves. We have had to sit in the middle of heated discussions in order to settle out matters of clearance and rights and responsibilities in the three-way relationships.

We found that lower school staff had to know beforehand the whys and wherefores of our coming to the schools. Inadvertent skipping along the hierarchical network meant recycling again to obtain renewed support from those above and below. Moreover, informal understandings or vague, general assent, either verbal or written, was found to be insufficient to the total task of clearance and reciprocal cooperation. We found that we had to sit down on a face-to-face basis with every affected staff member to explain fully our whys and wherefores of programming and evaluation. In other words, full consent and cooperation could not be obtained without each participant in the cooperative process understanding and taking a more or less personal, official commitment to the effort. The same discussion and justifications which applied to top staff had to be repeated within the total network of staff aiding this research. Absence of face-to-face explanations with key personnel at each level led to recontacting and starting all over again to insure the quality of work demanded for this study.

In retrospect, time consumed both in clearing channels, even when it meant shuttling between staff and offices, and in explaining and even meeting hard arguments with reasoned counter-arguments had led to effective cooperation with school staff. When we did not follow through on what is deemed as proper clearance, we have had to face the prospect of folding up the programming or the research effort. It also has meant that principal investigators and field workers have been frustrated many times by official and personal rebuffs because of our lack of understanding of organizational processes in proceeding to work with offices involving central and local school administrations. It has meant that we have had to confront supervisory personnel directly with hard arguments about why the work must proceed, even if their personal and official pride and functions have been questioned. It has meant that major supervisors in both research and school administrations have had to retract and modify as our total research needs and aims were balanced against what they have felt to be standing school policy and practice. In this continual give and take, we believe that we have reformed administrative orientations in both programming and value of research follow-up. We also believe that we have had to reform our own orientations towards schools and school administrators and take into account their own organizational burdens and hopes.

In retrospect for example, our selection procedure required staff to spend more time in testing their own sensibilities as well as scanning available test and grade profiles of students to determine who might be eligible. Our usual request for 30 to 50 students per school (2 to 3 times the number required to fill the spaces for the program from specific schools) encouraged staff to look farther than the few they already knew of as quite capable and talented. To "dip lower" than by previous standards and norms meant that the staff was being questioned about normal practice. Seven years later, skepticism about "lowering standards" to reach "hard core" youths no longer is met with practiced exclusion from the nomination lists. Instead, guidance and counseling personnel assigned to talent search are ready and willing to suggest names of troublesome but talented youths at any time, hoping that the program can work with the students to "set them straight".

b. Learning to Articulate Goals and Roles. As Eisenstadt suggests, those who are willing to reciprocate in institutional exchange may only be willing to do so when they can see that the new arrangements can "answer their need for some general stability or more specific needs and problems." The benefits gained from older arrangements must be evaluated against the benefits to be gained from adopting new procedures. The gains from the new must be high enough for voluntary compliance to new demands. In our case, the exchange must be of sufficient value that the school staff can accept new norms and added tasks in talent search, in selection, and in long-term evaluation and be able to replace older practices which may have had value also. As noted above, we had to learn to clarify what we wanted to do and to take into account the pervasibe personal and organizational concerns of the school staff in providing reasonable and informed justifications for doing more systematic evaluative work. What we describe here are the kinds of organizational burdens and constraints school staff did have to carry in acceding to our research requests. We also describe what have been and are some of the underlying goals and beliefs on our part which enabled us to explain and to justify our work to the school staff.

It must be remembered that our evaluation procedures placed extraordinary burdens on school staff. For one, research and teaching-guidance staff often felt that their own work was being evaluated. For another, we asked the staff to undertake new obligations out of the normal call of work. It was not that the staff were incapable of intensive talent search, incapable of accepting a selection procedure which takes away personal and professional discretion, and not able to accept the long-term burden of retrieving records of educational careers of students in the school district. They could do all of these tasks. It was that they had to do such tasks under organizational demands which either limited their efforts or else even rewarded their older practices in these matters. Changes in the scope and type of talent search, in the mode of selection, and in the long-term evaluation of treatment effects, all demanded at the same time, meant that we disrupted established organizational routine and added new ones.

(1) The Extraordinary Demand for Talent Search. We found very quickly that personnel assigned to talent search functions in more impacted schools were quite skeptical of our efforts to find talented youngsters who could be prepared for college. Our policy approach was to increase the skill levels and sensitivities of the regular school staff (usually guidance counselors) in scanning for talented youngsters in these schools. This meant a careful review of tests and grades; but it also meant use of observational data from the classrooms and the guidance offices. We encouraged the staffs to stretch the usual test and ability tracking norms, grades, and standards of school decorum.

The following chronicle, summarizing several rounds of discussions with fourteen school staff from the five districts, reflects the tenor of the arguments and counter-arguments.

Line Staff: "Well, where shall we establish a cutting point to recommend children for this program?"

College Representative: "Where do you normally 'cut' to recommend children for college preparatory work?"

Line Staff: "There are no college preparatory courses in junior high school although there are classes by different ability groupings."

College Representative: "What is the cutting point for assignment into these various classes?"

Staff: "Well, we usually recommend students who score _____ points and above."

Other Staff: "Just try finding students in our schools who score as high as the cutting point. I only know of student ____________________, ____________________, and ____________________."

Silence.

Staff: "Look, do you want us to nominate students from these schools and from these backgrounds? Even if their test scores are below the line? We don't do this in our system for our own programs."

Other Staff: "Why, in the later grades, we can't even ask these students to take college preparatory courses if they do not meet certain test and grade norms."

College Representative: "Well, we're willing to have these students nominated who may not meet school district norms but who are considered high potential on the basis of classroom behavior."

Staff: "Are you willing to take students who score _____, the average score, then?"

College Representative: "Yes, you can nominate students within this range if you and your staff feel that with some individualized help in our residential program during the summer, the student might 'turn the corner' in performance or motivation."

Staff: "With all this effort, wouldn't it be simpler to take only the bright ones who show no real problems or risks? You ask us to nominate more than you can take, but for all this work, you only take a certain number. Why not take the few bright ones? You and I know that these are the students who will work hard, be a credit to the school and their families, and get support from their families. Taking those 'near delinquents' means excluding these deserving children."

College Representative: "What we are seeking are students whose chances for going on in higher education are low unless information, skills, and resources are provided in one way or another. Children whose academic careers and family backgrounds are clearly headed towards college work will be carried by their families over the long run.

The students we are seeking are highly talented youngsters whose backgrounds are not that favorable without help."

"Given this interest, where and how will we find out whether such students can be helped or not?"

Staff: "But the children you have taken before are so 'average' and you never take all of those who are high achieving."

College Representative: "In relation to your judgments, how are our older SOP (Special Opportunities Program) children doing?"

One Staff Member: "Well, it is a mixed picture, but they certainly get wound up about the Oberlin program and keep asking when they can return to Oberlin."

Another Staff Member: "I'm not sure that the children you call 'average' are 'average'. How about student ___________ at your school, and student __________________?"

First Staff Member: "Come to think of it, they could be considered 'high average'."

Second Staff Member: "For these children, couldn't we recommend them to continue with summer work that could keep up their interest and their good work?"

First Staff Member: "But the way we recommend for these summer programs in the local schools is based on tests and grades as well as their ability grouping."

Second Staff Member: "Why not recommend them for these programs if they do show some ability or potential even if they do not meet test norms?"

Central Staff Member: "You know that test scores are approximate rather than exact. Test variability suggests that you could and should encourage these students if they are within an acceptable range. You should stand behind these recommendations if students are within the range of acceptable variation but below the suggested fixed score necessary for recommendation."

(2) Extraordinary Demand for Random Selection. With the additional request to allow an impersonal method of selection to determine who would be chosen for the special program, we faced serious objections until we understood the organizational and personal basis for such objections. The following list of factors appeared to constrain school staff to reject random selection as a preferred procedure.

(a) Professional Discretion. School staff typically believe that they are best able to identify children most likely to profit from a program. If researchers ask for the power to decide, the school staff may believe robbed of routine discretion in deciding who should or should not receive given educational services.

(b) Choice of "Players in School Contests." Any organization required the demonstration of "successful outcomes" to be able to motivate its staff and to channel organizational and societal rewards. Schools, like other "productive systems" tend to be evaluated by the "quality and quantity" of "output". Students become representative "players" for schools and their staff who are involved in personal and official "school contests" to demonstrate that they do "produce" excellence. Even colleges and universities use such laurels to argue that their school is excellent. Even if the best predictor of what comes out of the schools is what goes in as tested talent, the schools persist in claiming special gifts in training and producing such students.[5]

The "best organizational game" for a given staff member and his coalition is to propose their "best students" for special programs. In this way, they can guarantee positive outcomes, which in turn reflect well on themselves and their schools. Random, equi-probable selection, even if chosen by numbers in a bowl (actually employed in our research to encourage participation in selection in some schools), meant that staff could not guarantee that their best students would be placed. That this is a source of strain is shown by efforts to replace selected students with others who enjoyed staff preference, with appropriate comments about how much more deserving the non-selected was compared to the selected student. Arguments among staff and families of children who were pre-notified that they might be selected also reveal that certain students were already selected as ready to go. Subsequent arguments within the schools concerning the fact that some "high risk" students were selected in preference to the "more deserving" students revealed the intensity of the commitments to insuring "safe and valued outcomes" and in aiding those who already had worked very hard to "get ahead."

---

[5] At the college level, we have indications that the rate of "productivity" of undergraduate institutions is a function of the quality of "input" of able students. See A. W. Astin, "Productivity of Undergraduate Institutions," SCIENCE, 136 (1962), 129-135; "Undergraduate Institutions and the Production of Scientists," SCIENCE, 141 (1963), pp. 334-338.

(c) Rewarding Those Who Get Ahead. Given the ethic of social mobility and personal achievement, random selection violated societal norms about who should be rewarded to attend special programs. The very concept of viewing the special program as a reward for good and progressive behavior suggested that many staff members wanted more to reward than to find out whether the less motivated or troubled or seemingly less able could also be encouraged to achieve.

(d) "Incorrigibility" of the "More Needy." A belief in the seeming "incorrigibility" of the troubled students also affected staff willingness to search for talent and to support special programs to provide an extra boost. Given their range of experiences in the school and in personal efforts to aid the "hard core" student, the staff express much skepticism about doing anything significant with these youths. This belief is tied closely to the view that those who are getting ahead are more corrigible or responsive to extra boosts. In retrospect, we see that this belief is maintained when there is competition for rare and valued spaces in special programs. Under such limitations in special opportunities, the tendency seems to be to exclude those who they feel will require more assistance than most programs promise or deliver. In our seventh year of cooperative programming with the schools, the practiced exclusion of more "hard core" youths is no longer a serious obstacle to search for talent.

Countering Organization Constraints. Against this nexus of beliefs, practices, and rewards influencing search and selection, we based our arguments and discussion on the following principles.

(a) Commitment to Rational Knowledge. Whether in talent search, selection, or long-term evaluation, our arguments stressed staff commitment to rational and systematic knowledge in providing school services. This general commitment was reinforced in specific ways, as noted below.

(b) Commitment to Reducing Talent Loss and to Supporting Equal Opportunities. These two values are strongly held by school staffs. They remember those students whom they have been able to help through compensatory, special efforts. We argued from such personal knowledge of staff and our own growing evidence for alternative ways to reduce loss through talent search and evaluation of alternative programs.

(c) Commitment to Hard-Headed, Experimental Probes. School staff, having gone through educational methods courses where they learned the catechism of experimental research, were reminded of its value and this specific opportunity to complete a major study with their aid. This training and commitment has stood up well in our case.

(d) Commitment to and Delivery of Services. We clearly recognized that we were imposing additional burdens on time and resources by our requirements for talent search, random selection, and long-term evaluation. We stated that we stood ready at any time to absorb staff and fiscal costs of search, selection, and evaluative follow-up. Given such commitments, it was difficult for staff to reject our efforts on the basis of expense and cost alone. The basic issues turned more on implementing our instructions in ways that were clear and uniform. We were prepared to provide necessary staff and time of our own as well as to absorb the charges which accrue to school staff in these interconnected research tasks.

(e) Commitment to Long-Term Support and Research/Policy Reciprocities. There is a widespread belief, subscribed to today with increasing vigor, that social scientists research "disadvantaged" populations for their own gains, without reciprocal returns to either the public agencies or their clients. Staff clearly know that they have been subject to one-shot case studies in which the schools and their constituencies have been the basis of scholarly and graduate student mobility. As one staff member said, research studies "benefit the researchers and their graduate students, and the poor children receive nothing in return for tons of research studies."

This theme was dunned into the ears of the researchers well before minority populations began to press for reciprocal benefits. Our own policy commitment has been very clear. As researchers we are committed to reciprocal benefits of policy research. In major and minor ways, we have demonstrated our ability to deliver on reciprocal benefits at both the college and the pre-college levels. We have been involved actively in the continual refinements in programming on the basis of feedback of information from field visits and studies. We have even had to ask ourselves whether the original research design might become unworkable as we participated in refinements in programming.[6] We believe that our research shows that sound, basic research can also be of benefit to the agency and its clients.

(f) Commitment to Values of Fairness and Equity in Selection and Treatment. Given that special, college-preparatory programs involving colleges and the public schools are still rare, we got enbroiled immediately in the problem of who decides on selection and who should be educated. We have noted above that staffs did not want to make their schools look "bad" in a prestige rating involving their children as players for the school team. But neither did they wish to exclude more "hard core youths" who they knew could profit from the extra, compensatory boost. It was the scarcity of special opportunities that made them push towards rewarding the more deserving.

Given such scarcity, we argued that we truly wanted to know whether both the well-motivated students and the more needy ones could benefit from a special program. It is clear that the rule of random selection broke the impasse in such instances. It was fair when spaces were too few, it was equitable in allowing knowledge to grow without excluding either the more achieving students or those with latent talent.[7]

---

[6] 6. See Gordon Welty, "Experimental Design and Applied Research," CALIFORNIA JOURNAL OF EDUCATIONAL RESEARCH, 20 (1969), pp. 40-44. Welty explores the measurability of continuing changes in programming in affecting the interpretability of results.

In some instances staff pre-selected and pre-notified students and their families (and the associated coalition of school staff) because they wanted the school to "win" and because they wanted to please their most active and vocal constituencies. We reduced the political risk of displeasing those not selected by allowing the staff to state that they honestly gave each child an equal and fair chance to enter the program. Given the few spaces and the pressing need to develop a program involving the "hard core", few parents and staff could object. After some experience, staff looked forward to pulling slips out of a bowl or allowing our staff to randomly select the students. We always had "special requests" also, which suggests that without expanding the range and type of special programs or else making such opportunities a normal part of any public school, we will face enormous pressures to reward only the most able students.

We constantly had to cope with the question of equity with regard to whether the control group children deserved anything for being subjects for study. We argued that if the program worked, then the control group children would be disadvantaged. If serious, disruptive effects on the life of the child in the program occurred, however, then the control group child would be advantaged.[8]

---

[7] We have now available in the literature ready means for systematic evaluation of the power of alternative designs in controlling plausible rival hypotheses to the research hypothesis. The terminology, labeling, and numbering of alternative designs and the listing of the sources of invalidity are fully developed by Donald T. Campbell and Julian L. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," in Nathan L. Gage, ed., HANDBOOK OF RESEARCH IN TEACHING (Chicago: Rand MacNally and Co., 1963), pp. 171-246. Campbell and Stanley's discussion also provides guides to the problem of identifying organizational and extraorganizational constraints and encouragements to given arrangements and design procedures. See also Campbell's "Reforms as Experiments," op. cit., passim. See also D. T. Campbell, "Administrative Experimentation, Institutional Records, and Non-Reactive Measures," J. C. Stanley, IMPROVING EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS (Chicago: Rand MacNally, 1967), pp. 257-291.

In retrospect, on the principle of reciprocal benefits and commitment to hard-headed research, we should have followed the advice of a tough critic of our research design in a major school district. That staff member pressed for some return to the control school children directly over time rather than a delayed contribution after the completion of the study itself. Future research studies involving families from more economically impoverished communities might note that (1) there are more children who require servicing than a special program can ever take; (2) that the rate of response to interview and mailed probes is lower than in most longitudinal, panel studies of student plans and achievements; cooperation must be encouraged from control groups especially.

Any research design can provide some limited service, which is both theoretically and policy relevant for the control group(s) involved. Multi-level treatments, from the most comprehensive to the minimal which might still make a difference in educational plans and achievements, can be mounted so all participants in a study can receive some benefits, (if there are benefits). In our design, if we had provided some annual service involving the mailing of career

---

8. This concern about the "negative consequences" of short-term efforts which disrupt existing arrangements are expressed by both highly supportive and nonsupportive individuals and groups. The overriding effects of "too much" for children who are used to "too little" without adequate instruction into the means and rules for mobility seem to predominate in these concerns.

planning information to the control group children as well, with the proviso that they will keep in touch with us, we might have been able to maintain a rate or return on evaluative follow-up as high as that of our students in the educational program itself. Such limited services might be sufficient to permit "control group" families to prepare further and to place their children into valued educational arrangements after high school. At the present time, this approach seems both fairest in terms of the pressing need to maintain the efforts of such youths in their schooling and to obtain hard information on the level of intervention necessary to bring all talented youths up to their potential in our society.

Tracing Educational Careers Over the Long Term. We have demanded extraordinary services from the public schools in tracing the educational careers of the students in the study. We work at the edges of some difficult legal, political, and moral issues about data access and the uses of such data. Moreover, we work on the edges of concerns about informed consent in participation in studies and in long-term evaluative follow-up.[9] We are prepared to explain our work fully, as we have had to do with school staff and the parents involved,[10] in order that basic studies which also are relevant to policy can proceed on an open and informed basis.

---

[9] A major set of concerns which involve social science research efforts at systematic procedures for assignment-selection, treatment, and assessment of administered interventions involve both organizational and extraorganization constraints. Among other considerations are (1) the disruption of organizational integrity, (2) of unfair treatment or of lack of equal opportunity to receive rare and novel services, (3) of the protection of the privacy and the minimization of disruptive consequences of given interventions and then assessment. These perceived consequences which directly or indirectly touch the participants in the administration and the receipt of public services represent serious challenges to the development of feasible, legal, and moral arrangements in research design and assessment. (continued)

---

[9] This concern appears in part in the "Important Clarification of the Public Health Service Policies Relating to Research on Human Subjects," AMERICAN SOCIOLOGIST, 2 (February, 1966), on the United States Public Health Service guidelines concerning experiments with human subjects. John Lear's "Do We Need New Rules for Experiments with People?" SATURDAY REVIEW, February 5, 1966, pp. 61-70 explores some of these concerns in the biomedical field. Continuing exchanged in Letters to the Editor in SCIENCE on experimentation and privacy and subject's consent indicates that we may have to formalize many arrangements which are made by trial and error without full assistance from participants and without legal advice. J. C. Maloney, "Psychological Experiments without Subject's Consent," SCIENCE 152 (June 10, 1966), p.1455; John W. Hamblen, "Preservation of Privacy in Testing," SCIENCE 151 (March 11, 1966), p. 1174 in response to Dale Wollfle's editorial on "Psychological Testing and the Invasion of Privacy," (December 31, 1965) SCIENCE.

Wolf Wolfensberger's "Ethical Issues in Research with Human Subjects" in SCIENCE, 155 (January 6, 1967), pp. 47-51 suggests some guidelines for establishing arrangements and how consent arrangements can be gauged relative to likely consequences. Letters to the Editor in the March 31, 1967, issue of SCIENCE, pp. 1617-1618, further carry on the discussion and debate.

See also Lee Rainwater and David J. Pittman, "Ethical Problems in Studying a Politically Sensitive and Deviant Community," SOCIAL PROBLEMS 14 (Spring, 1967), pp.357-65; Kai T. Erickson, "A Comment on Disguised Observation in Sociology," SOCIAL PROBLEMS 14 (Spring, 1967), p366-72; and John Walsh, "Antipoverty R & D: Chicago Debacle Suggest Pitfalls Facing OEO," SCIENCE (September 19, 1969).

[10] This discussion has not detailed the questions involving the appropriate legal and moral structures involving access to public records and the ties between client (pupils and their legal guardians) and the researchers. This concern is a topic for discussion in itself. See Frederick H. Treash, "Do Files Invade Privacy?" ELYRIA CHRONCILE TELEGRAM, Thursday, April 2, 1970, p. 32. For a discussion of the nexus of issues, see Donald T. Campbell, "Prospective: Artifact and Control," in R. Rosenthal and R. L. Rosnow, eds., ARTIFACT IN BEHAVIORAL RESEARCH, (New York: Academic Press, 1969), esp. pp. 367-372. On data bank concerns, see J. Sawyer and H. Schecter, "Computers, Privacy, and the National Data Center: The Responsibility of Social Scientists, " AMERICAN PSYCHOLOGIST 23 (1968), pp. 810-818. On the procedure to insure privacy through "mutually insulated data banks," see R. D. Schwartz and S. Orleans, "On Legal Sanctions," UNIVERSITY OF CHICAGO LAW REVIEW 34 (1967), pp. 274-300. These matters will be explored in detail for this study in a later publication.

We also have had to work patiently with school officers (as they have had to extend their patience with us) to develop administrative and staffing arrangements for tracing educational careers. In general, the major burden has fallen on our own research staff to retrieve the necessary data, but this achievement in itself suggests that the school district personnel do want to know whether our form of cooperative educational program can help them and the children in the impacted public schools.

Conclusion

Social scientist typically get excellent book learning in the matter of experimental methodology but little in the way of systematic guidance and instruction in mounting such hard-headed research designs in field settings. Little is transmitted regarding what is involved in retaining effective legal and moral ties with those parties who can provide hard and significant longitudinal time-series data of value to policy and to basic study. We have sought to fill this lack with both a theoretical perspective or organizational analysis on processes associated with role crystallization, as well as an interpretive description of how we sought to clarify and justify our work to those who aided us in completing the research.

It is clear that social scientists interested in field studies must become familiar with ways to establish "status equivalent" contacts and commitments to obtain organizational consent. It also is clear that they must become adept at sensing organizational constraints and encouragements to experimentally grounded evaluation studies. Moreover, those who desire firm, interpretable evidence from field settings must be able to provide the kind of legal, moral, and political justifications which can serve to legitimate honest and experimental and quasi-experimental research which can counter-balance constraining organizational and extraorganizational influences. Finally, such students of policy and basic research must be able to create role relationships in the interstitial areas involving members of two organizations, such that fair and equitable exchanges can occur as researchers evaluate public policy. Public servants, their clients, and their constituencies can then obtain the kind of hard-headed assessments necessary to settle some difficult issues involving effective public services. We suggest some principles and justifications which are appropriate to educational policy work. We believe that such principles and arguments are applicable with modification to many other institutional agencies in housing, family welfare, police protection, and related public and community services.

B I: SUGGESTIONS FOR SELECTION PROCEDURES FOR SUMMER SPECIAL EDUCATIONAL OPPORTUNITIES PROGRAM

Aims of Selection: A major goal in the suggestions listed below for selecting 7th and 8th graders for the summer enrichment program at Oberlin College is to recruit students with high potential and ability for work in higher education. The program is experimental in the sense that we need to know which type of student will benefit and in what ways they will benefit from the planned program of instruction and co-curricular activities. Thus, the hope is to recruit both "low risk high achievers" who show up in many ways as able students and some "high risk high achievers" who show up in some ways as talented students. Listed below are selection procedures and the criteria on which they will be based. In addition, suggested proportions of "low risk-high promise" students and "high risk-high promise students" are described.

Criteria: Three sets of criteria are important in selecting students for this program: (1) Criteria which define "eligibility" among students from "severely disadvantaged backgrounds"; (2) Criteria which define "promising talent," both as to "risk" and "promise" for higher education; and (3) Criteria which define "behavior problems" which may be critical for the success and effectiveness of this program.

(1) "Severely Disadvantaged Background": Three criteria jointly define the eligibility of students for this program.

(a) Limited Financial Resources. To be eligible, the student comes from a family with limited financial resources. A student may be considered eligible if his family would have to cut down on the purchase of basic necessities (clothing, food, shelter) in order to help finance his future education.

(b) Limited Neighborhood Opportunities. The student also lives in an area where neighborhood resources are less developed in encouraging and stimulating higher educational plans and aspirations.

(c) Minority Status. Because of race, ethnicity, or regional background, the students are not likely to receive the same opportunities for higher education as those who are not subject to discrimination.

(2) "Promising Talent": Among 7th and 8th graders from given schools in these neighborhoods, the following criteria are proposed for selection of students on the basis of "promise or talent" and "risk." We can consider all students who show up well in one or more of these criteria as "promising" or "talented." On the other hand, we feel (whether correctly or not) that there is less risk with some students and more with others among the "talented." One way of identifying the risk involved in aiding a given student is on the basis of whether he shows up well in all of the criteria of selection for talent or he shows up with only some positive signs. Those who show up with many positive signs would be considered "low risk" as against those who show up with only one positive sign (teacher nominations, for example) among all potential criteria.

Four sources of "talent" are recommended to select the students:

(a) Information from a test(s) of intelligence.

(b) Information from achievement tests.

(c) Information from grades in school.

(d) Information from teacher nominations.

The "cutting points" for each selection criterion will have to be worked out with each school system because different test measures are employed among the participating groups. If the "normal" cutting points for identifying "good risks" for higher education are established for the schools, the "high promise" students would represent those who are above that cutting point, the "low promise" students would be those who are below that cutting point. It should be noted that for each of the selection criteria, all of the students in the given grades may or may not be of "high promise." They may be of "high promise" in some criteria but of "low promise" in others. What is desired here is as large a "talent pool" by available criteria as possible to give these students a chance to participate in the summer program.

(a) Intelligence Test Scores: If the normal cutting point is 130 points, the lower limit would be at 110 points and above. The general rule would be to include in this listing those students who are about 20% below the normal cutting point for selection into a college preparatory program.

(b) Achievement Tests: The lower limit would be one grade score below the given grade in which the student is found (in academic-college prep work).

(c) Grades: The lower limit would be one grade below the given grade (averaged) which is considered necessary for counseling students to follow through for college work in academic-college preparatory subjects.

(d) Teacher nominations: The procedure to follow would be to suggest any other students who would not be "high promise" students by test or grade scores as well as those who are high achievement by other criteria. This is to insure that as large a number of "talented students" by any of these four measures (a-d) will appear on the eligible list.

(3) "Behavioral Problems": This is to aid in the selection of students for the summer program so that a wide range of talented students can be aided. The intent here is to note any serious problems which students might bring to the residential experience which might be detrimental to themselves and to other students in the program. The aim is not to exclude the students with behavior problems, but to include those who appear to be talented and who could benefit from this program.

This summer program, as structured now, cannot adequately service the talented students who have severe behavioral problems in the schools and the community. These students may require more support and assistance than this program can provide. The point to note is where school officials (teachers, counselors and principals included) can provide information on students as to severity of behavioral problems, the staff of this program can pre-plan for these students. We would like to see a few students (up to 10%) who are of high promise but have some behavioral problems in order to aid these students where possible.

(4) Procedure: The following is a set of suggestions to enable a fair selection of students from different talent-risk groupings in relation to the goals of the experimental program at Oberlin.

(a) Develop a list of nominees in each of the four criteria of talent in terms of "high achievers" or those above the selected cutting point and "low achievers" or those below the cutting point.

1. Begin with the list of students on the basis of intelligence tests.
2. Next, develop a list of students on the basis of achievement tests.
3. Next, develop a list of students on the basis of grades. From this list, exclude all students below the cutting point (high risk students) in grades, _if_ they do not show up as promising in intelligence or achievement scores.
4. Finally, develop a list of promising students with nominations from teachers in the 7th and 8th grades. This list should contain those names of students _who do not show up in any or all of the criteria above_ as well as those who are of "known promise" by one or more of the first criteria.
5. Separate the boys' list from the girls'.

(b) Classify each student into one of two "promise-risk" groups below.

1. Students who show up as "high promise" students in _all four_ of the talent criteria above should be placed in the _High Achievers_ group.
2. If the size of the group which is of "high promise" in all four criteria is small, place the students who are of high promise in at least three of the four criteria in the _High Achieving_ group. If not, place them in the _Low Achieving_ group below.
3. Place the students who are high achieving in only two or one of the four criteria of talent into the _Low Achieving_ group.

(c) What is desired is a large proportion of "High Achievers" (80%) and a lesser proportion of "Low Achievers" (20%).

1. Rate each student as to severity of behavior problems for the summer program. Exclude those students who require a level of supervision which cannot be provided by the program.

2. Sample within the "High Achievers" group until the number of students desired is obtained. (80% of the total number of students coming to the summer program from the schools.)

3. Sample within the "Low Achievers" group until the number of students (20% of the total number of students invited from the schools) is obtained.

4. Go over the list to see if the percentage of students with somewhat severe behavior problems represents no more than 10% of the total number of students. Among this 10% the hope is to obtain students who are "natural leaders" who may be somewhat difficult to handle but are of high potential.

B II: OTHER INSTRUMENTS AND DOCUMENTS

The following are not included in this report because of their bulk or because of their limited interest. Persons who wish to examine them are welcome to write to us for copies.

1. Pupil Data Forms
2. Interview Schedules and Record Forms
3. Letters requesting cooperation from schools, parents, and other sources.
4. Mail questionnaires to families for follow-up information.